UNIVERSIDADE DE LISBOA

Faculdade de Farmácia

Antibody engineering for HIV therapeutics: From fusion inhibition to delivery of genetic regulators

Ana Catarina Cunha Santos

Orientador: Professor Doutor João Manuel Braz Gonçalves

Tese especialmente elaborada para obtenção do grau de Doutor em Farmácia, especialidade Biotecnologia Farmacêutica

2017

UNIVERSIDADE DE LISBOA

Faculdade de Farmácia

Antibody engineering for HIV therapeutics: From fusion inhibition to delivery of genetic regulators

Ana Catarina Cunha Santos

Orientador: Professor Doutor João Manuel Braz Gonçalves

Tese especialmente elaborada para obtenção do grau de Doutor em Farmácia, especialidade Biotecnologia Farmacêutica

Júri:
Presidente: Doutora Matilde da Luz dos Santos Duque da Fonseca e Castro, Professora Catedrática e Diretora da Faculdade de Farmácia da Universidade de Lisboa
Vogais:
- Doutora Isabel Maria Corrêa Calvente Barahona, Professora Associada
Instituto Superior de Ciências da Saúde Egas Moniz;
- Doutora Ana Lúcia Freitas de Mesquita Barbas, Investigadora
Instituto de Biologia Experimental e Tecnológica, Responsável pelo *Protein Production Platform Laboratory*;
- Doutor Frederico Nuno Castanheira Aires da Silva, Investigador Auxiliar
Faculdade de Medicina Veterinária da Universidade de Lisboa;
- Doutor José António Frazão Moniz Pereira, Professor Catedrático
Faculdade de Farmácia da Universidade de Lisboa;
- Doutor João Manuel Braz Gonçalves, Professor Associado com Agregação
Faculdade de Farmácia da Universidade de Lisboa, Orientador.

Fundação para a Ciência e Tecnologia – Ministério da Educação e Ciência:
SFRH/BD/73838/2010

2017

Ana Catarina Cunha Santos was financially supported by a PhD fellowship (SFRH/BD/73838/2010) from Fundação para a Ciência e a Tecnologia—Ministério da Educação e Ciência (FCT-MEC), Portugal.

The research described in the present thesis was performed from September 2011 until February 2017 under the supervision of Professor Doutor João Gonçalves at the URIA-Unidade de Retrovírus e Infeções Associadas and Research Institute for Medicines (iMed.ULisboa) of the Faculty of Pharmacy, University of Lisbon, Lisbon, Portugal. This project was also performed in collaboration with the HIV Evolution, Epidemiology and Prevention and Host-Pathogen Interactions groups from iMed.ULisboa, Faculty of Pharmacy, University of Lisbon, the Physical Biochemistry of Drugs & Targets group from Instituto de Medicina Molecular, School of Medicine, University of Lisbon, the Animal Nutrition and Biotechnology group from Interdisciplinary Centre of Research in Animal Health (CIISA), Faculty of Veterinary Medicine, University of Lisbon and the TechnoAntibodies (TechnoPhage S. A.), Lisbon, Portugal.

The results described in this thesis were included in manuscripts already published or in preparation, in peer-reviewed journals and posters of international and national scientific meetings:

Cunha-Santos C, and Joao Goncalves. CXCR4-targeted nanobody for delivery of modulators of HIV gene expression. Manuscript in Preparation.

Cunha-Santos C, Figueira TN, Borrego P, Oliveira SS, Couto A, Cantante C, Santos-Costa Q, Azevedo-Pereira JM, Fontes CM, Taveira N, Aires-Da-Silva F, Castanho MA, Veiga AS, and Goncalves J. (2016). Development of synthetic light-chain antibodies as novel and potent HIV fusion inhibitors. AIDS. 30(11):1691-701. doi: 10.1097/QAD.0000000000001108.

Cunha-Santos C, and Goncalves J. CXCR4-Targeted Delivery of Anti-HIV Small Interfering RNA. 8th iMed.ULisboa Postgraduate Students Meeting 14-15 July 2016, Lisbon, Portugal.

Cunha-Santos C, and Gonçalves J. Nanobody-Based Specific Delivery of Therapeutic Molecules Against HIV Infection. 7th Annual PEGS EUROPE - Protein & Antibody Engineering Summit 02-6 November 2015, Lisbon, Portugal.

Cunha-Santos C, and Gonçalves J. Nanobody-Based Delivery of Therapeutic molecules against HIV infection. 7th iMed.ULisboa Postgraduate Students Meeting 15-16 July 2015, Lisbon, Portugal.

Figueira TN*, Cunha Santos AC*, Borrego P, Fontes C, Taveira N, Aires da Silva F, Goncalves J, Castanho MARB, and Veiga AS. Protein-lipid membrane interactions of a broad and potent engineered single domain antibody against HIV. IX PhD Students Meeting 25-27 March 2015, Lisbon, Portugal.

* These authors contributed equally to the work.

Cunha Santos AC*, Figueira TN*, Oliveira S, Borrego P, Couto A, Cantante C, Costa Q, Veiga AS, Pereira JM, Fontes C, Taveira N, Aires da Silva F, Castanho, MARB, and Goncalves J. Engineering synthetic domain antibodies with broad and potent HIV-1 neutralization, by direct binding to gp41 viral glycoprotein. 6th iMed.ULisboa Postgraduate Students Meeting 2nd July 2014, Lisbon, Portugal.

* These authors contributed equally to the work.

Santos CA, Oliveira S, Borrego P, Couto A, Cantante C, Costa Q, Pereira JM, Fontes C, Aires da Silva F, Taveira N and Gonçalves J. Engineering synthetic domain antibodies with broad and potent HIV-1 neutralization, by direct binding to gp41 viral glycoprotein. 5th iMed.UL Postgraduate Students Meeting 18th July 2013, Lisbon, Portugal.

Santos C, Oliveira S, Borrego P, Couto A, Cantante C, Costa Q, Pereira JM, Fontes C, Aires da Silva F, Taveira N and Gonçalves J. Engineering synthetic domain antibodies with broad and potent HIV-1 neutralization, by direct binding to gp41 viral glycoprotein. 8th Annual PEGS – the Essential Protein Engineering Summit 30-4 May 2012, Boston, MA, USA.

Santos C, Oliveira S, Couto A, Costa Q, Pereira JM, Fontes C, Aires da Silva F, and Gonçalves J. Design of synthetic single domain VL antibodies that inhibit HIV-1 infection by direct binding to gp41 HR1. HIV-AIDS Conference FLAD 7-9 September 2011, Lisboa, Portugal.

Santos C, Costa Q, Couto A, Pereira JM, and Gonçalves J. Design of synthetic single domain VL antibodies that inhibit HIV-1 infection through binding to gp41 HR1. EUROPRISE Network of Excellence annual conference "Rational Design of HIV Vaccines and Microbicides" 17-19 November 2010, Lisboa, Portugal.

Santos C, Couto A, and Gonçalves, J. Design of synthetic single domain VL antibodies that inhibit HIV-1 infection through binding to gp41 ectodomain. MicroBiotec 28-30 November 2009, Vilamoura, Algarve, Portugal.

The study described herein also contributed for the publication of manuscripts in distinct areas from the present thesis:

Gouveia Z, Carlos AR, Yuan X, Aires-da-Silva F, Stocker R, Maghzal GJ, Leal SS, Gomes CM, Todorovic S, Iranzo O, Ramos S, Santos AC, Hamza I, Gonçalves J, and Soares MP. Characterization of Plasma Labile Heme in Hemolytic Conditions. (submitted)

Figueira T, Freire J, Cunha-Santos C, Heras M, Gonçalves J, Moscona A, Porotto M, Veiga AS, and Castanho M. (2017). Quantitative analysis of molecular partition towards lipid membranes using surface plasmon resonance. Scientific Reports. 7:45647. doi: 10.1038/srep45647.".

AGRADECIMENTOS

Gostaria de agradecer em primeiro lugar ao meu orientador Professor Doutor João Gonçalves. Obrigada por ter acreditado em mim e no meu potencial e por me ter acompanhado ao longo deste caminho difícil. Obrigada também ao Professor Doutor José Moniz Pereira por me ter aceite no seu departamento e sempre me ter apoiado.

Gostaria de agradecer também a todos os meus colegas que durante todos estes anos me ajudaram a nunca desistir e dar sempre o melhor de mim. Obrigada ao Pedro, Cátia, Ana Catarina Santos (minha homónima) e Soraia Oliveira pela vossa companhia e apoio ao longos dos anos. Obrigada à Mariana Santa-Marta e Paula Brito pelos vossos conselhos sábios e paciência. À Andreia Couto, obrigada pela orientação e apoio. Obrigada ao Acil pela ajuda no trabalho, discussões científicas e humor estrondoso. Aos meus restantes colegas Lídia Fonseca, Inês Soeiro, Luís Ferreira, Sofia Romano, Vasco Gonçalves, Cheila Rocha, Joana Oliveira, Nuno Saraiva, Íris Couto, Frederico Aires da Silva, Sara Maia, Sofia Narciso, Sofia Cerqueira, André Ramos, Carina Pereira, Rita Nogueira, Galber e Renato Alves obrigada pelo vosso apoio e partilha.

Um obrigada muito especial ao laboratório do Professor Nuno Taveira, em especial aos meus queridos colaboradores Pedro Borrego, Francisco Martin e Cheila Rocha. Obrigada pela vossa contribuição valiosa para a concretização desta enorme tarefa. Muito obrigada também ao laboratório da Professora Cecília Rodrigues pela vossa ajuda, compreensão e empréstimo de reagentes e materiais, em especial ao Pedro Rodrigues e Pedro Dionísio pela vossa boa disposição e ao André Simões e Pedro Borralho pela vossa ajuda e compreensão. Obrigada também à Maria João e Elsa Rodrigues pela ajuda e discussão científica. Obrigada ao grupo do Professor José Miguel Pereira, em especial à Quirina Costa, pela vossa contribuição e ajuda. Um muito obrigado ao laboratório do Professor Miguel Castanho, em especial ao Tiago Figueira e à Salomé Veiga. Muito obrigada pela vossa ajuda e apoio incondicional. Obrigada Tiago pela tua disponibilidade e profissionalismo. Obrigada ao laboratório da Professora Madalena Pimentel pela partilha e ajuda. Obrigada também ao laboratório e pessoa do Professor Carlos São José, Catarina Batista e Nuno Santos pela vossa colaboração e ajuda. Obrigada Sofia Fernandes pela tua amizade.

Um agradecimento muito especial ao laboratório e pessoa do Professor Carlos Fontes na FMV da Universidade de Lisboa. Por todo o apoio, ajuda e ensinamento. Por me ter aceite no seu excelente laboratório e me ter proporcionado todas as condições e conhecimento para o avanço do meu trabalho. Em especial à Ana Luís (obrigada pela amizade e companheirismo) e Vânia Fernandes (obrigada pela tua preciosa ajuda). E claro ao Jorge, obrigada pelo teu contributo e partilha de conhecimento.

Obrigada a todos os meus amigos por toda a compreensão e carinho e um obrigada muito especial ao João por me ter acompanhado em parte desta caminhada até os nossos caminhos se separarem.

Obrigada à minha família. Um especial aos meus tios António e Alzira, por tudo o que me têm ajudado desde que me mudei para a capital.

Obrigada à minha família "adotada" por todo o apoio, ajuda e carinho.

Aos meus pais. Obrigada por todo o sacrifício, amor e apoio incondicional e sobretudo pela vossa paciência para comigo nesta longa jornada. Vocês são a minha inspiração…

À minha mana, obrigada por estares lá sempre para mim… Ao meus cunhados pela paciência e compreensão. À minha pequenina sobrinha Matilde por tornar os meus dias mais risonhos.

A ti Pedro, obrigada… Por toda a paciência, apoio, ajuda, espírito crítico e companheirismo. Obrigada por sempre acreditares em mim, mesmo nos momentos mais difíceis…. Obrigada por estares sempre ao meu lado. Conseguimos…

E finalmente aos meus avós por toda uma vida de amor incondicional. Gostaria muito que estivessem cá para ver… Esta tese é para vocês…

ABSTRACT

Despite latest advances on antiretroviral therapeutics, HIV infection continues to be a chronic condition with serious complications and burden costs of treatment. The rapidly emergence of resistant strains constitutes a major contributor for failure of standard therapy. Accordingly, there is a continuous need for design of innovative HIV inhibitors against conserved and important targets on virus replication cycle. On the other hand, the exploration of novel classes of HIV inhibitors results in a necessity for novel efficient and specific delivery systems.

In the present thesis, we manage to target HIV infection through two distinct therapeutic approaches based on single-domain antibody (sdAb) technology. In the first one, we developed a novel fusion inhibitor of HIV by rationally engineering the binding sites of a stable light-chain scaffold from rabbit-origin. From five sdAbs with specificity to the conserved and crucial-to-fusion N36 viral sequence, one potently inhibited HIV-1 infectivity. The selected fusion inhibitor, named F63, was capable of broadly inhibiting distinct strains of HIV-1 with an antiviral potency similar to the T-20 entry inhibitor, clinically approved for HIV/AIDS treatment. Moreover, F63 presented an effective neutralization activity against HIV type 2, which constitutes an advantage relative to the limited antiviral breadth of T-20. In the second approach, we took advantage of a previously validated variable domain from camel (nanobody) to design a strategy for CXCR4-targeted delivery of HIV expression inhibitors. The engineered nanobody-based vehicle delivered an anti-HIV small interfering RNA specifically to CXCR4-bearing cells—susceptible to HIV infection—being the activity of the delivered RNAi effector proved by virus promoter silencing and inhibition of HIV replication. A second version of this construct allowed to deliver a zinc-finger transcription factor, which functionality depends on CXCR4-mediated endocytosis as mechanism of cellular crossing-over.

In resume, this study demonstrates the potential of engineering single-domain antibodies for HIV therapeutics and gives insights into the design of novel antiviral drugs.

Keywords: Single-domain antibody; HIV, Fusion inhibitor; Targeted delivery; Nanobody.

RESUMO

O Síndrome da Imunodeficiência Adquirida (SIDA) apresenta-se como um dos principais problemas de Saúde Pública a nível mundial, tendo sido responsável por milhões de vítimas desde a sua descoberta. Esta doença é caracterizada por uma grave deterioração do sistema imunológico dos indivíduos afetados, o que resulta no aparecimento de infeções oportunistas fatais. A SIDA foi pela primeira vez descrita em 1981, em consequência da deteção em indivíduos jovens de doenças resultantes de uma imunodeficiência demarcada. Estes acontecimentos culminaram no isolamento dos agentes etiológicos da doença, o Vírus da Imunodeficiência Humana tipo 1 e tipo 2 (HIV-1 e HIV-2 do inglês *Human Immunodeficiency Virus*). A síndrome causada pelo HIV-1 é mais severa e a taxa de transmissão apresentada por este tipo de vírus mais elevada. Devido a estas características, o HIV-1 constitui um alvo terapêutico preferencial relativamente ao HIV tipo 2. Apesar dos desenvolvimentos recentes na terapia antirretroviral, uma estratégia eficaz contra a infeção por HIV continua a estar indisponível, sendo o contínuo aparecimento de estirpes resistentes um dos principais obstáculos para a erradicação desta doença. Como tal, o desenvolvimento de novas estratégias terapêuticas assim como de métodos para a sua entrega eficiente continua a constituir uma necessidade urgente na terapia desta grave doença.

A engenharia de anticorpos surge como uma ferramenta importante para o desenvolvimento de novas terapias biológicas. Esta abordagem é particularmente importante no contexto da infeção por HIV devido à dificuldade em isolar anticorpos neutralizantes e com largo espectro de ação a partir de indivíduos infetados. Este facto prende-se com os diversos mecanismos de evasão imunológica apresentados pela glicoproteína viral de superfície responsável pelo processo de entrada do vírus infecioso na célula hospedeira. Entre estes destaca-se a localização de difícil acesso de sequências/regiões virais conservadas e críticas para o ciclo replicativo do vírus, que impossibilita o seu reconhecimento por anticorpos de tamanho tradicional. Neste contexto, a redução do tamanho do anticorpo para formatos pequenos, como os anticorpos de domínio único, constitui uma vantagem tremenda na ligação a estes epítopos. Os anticorpos de domínio único são apenas constituídos pelo domínio variável da cadeia leve ou pesada de um anticorpo. Estes fragmentos podem ser obtidos a partir de outros formatos de anticorpos ou selecionados a partir de bibliotecas. De entre estas, as

bibliotecas sintéticas são as que permitem maior controlo sobre a sua representatividade, visto que a sua diversidade pode ser cuidadosamente estudada e inserida por engenharia genética. Por outro lado, certos anticorpos de tamanho único, como os "nanobodies", permitem o uso de alvos celulares com objetivos terapêuticos devido não só ao seu tamanho reduzido, mas a propriedades favoráveis de ligação. Adicionalmente, estes fragmentos de anticorpos possuem extraordinárias características biofísicas que os tornam atrativos para diversas aplicações, nomeadamente engenharia genética para fins terapêuticos. Estes formatos em particular permitem a exploração de novas abordagens em medicina, como a entrega específica de moléculas terapêuticas.

A glicoproteína Env é responsável por mediar o processo de fusão entre o vírus de HIV e a célula hospedeira, apresentando-se como um alvo terapêutico atrativo devido à sua acessibilidade na superfície da partícula viral e papel fundamental no processo de entrada. Como tal, o Capítulo II desta dissertação de doutoramento teve como objetivo desenvolver um novo inibidor de fusão da infeção por HIV através da ligação de um anticorpo de domínio único a uma zona conservada e crítica da glicoproteína Env—região N36. Para tal e baseados numa estratégia racional de desenho, construímos uma biblioteca sintética de anticorpos de domínio único usando um fragmento variável de coelho como base. As regiões de ligação ao antigénio deste fragmento foram ampliadas e modificadas de modo a tornarem-se mais flexíveis e aptas para ligar à zona extremamente côncava onde a sequência N36 se situa. A partir da biblioteca sintética foram selecionados, pela tecnologia de "phage display", cinco fragmentos de anticorpos distintos que apresentaram ligação específica ao péptido N36 e impediram a infeção pela estirpe laboratorial HIV-1 NL4-3. Um destes fragmentos, F63, apresentou uma capacidade de neutralização potente e abrangente de diversas estirpes virais primárias quer de HIV-1 como de HIV-2, em alguns casos superior ao único inibidor de fusão aprovado no mercado, T-20. Este facto poderá em parte ser explicado pelo epítopo do fragmento F63 apenas parcialmente se sobrepor à região de ligação do péptido T-20. O anticorpo de domínio único F63 apresentou ainda uma ligação específica a membranas lipídicas, uma propriedade previamente associada a outros potentes inibidores de entrada.

A exploração de novas classes de inibidores antirretrovirais resulta numa necessidade crescente de desenvolver plataformas e estratégias para a sua entrega específica. Como tal, o Capítulo III da presente dissertação de doutoramento teve como objetivo a entrega específica de inibidores da expressão viral a células suscetíveis à

infeção por HIV. Para tal, tomámos partido de um domínio variável de camelo ("nanobody") previamente validado para o co-recetor de HIV-1 CXCR4 e construímos uma quimera como veículo de entrega de uma pequena molécula terapêutica de RNA de interferência (siRNA do inglês *small interfering RNA*). Os "nanobodies" são domínios variáveis autónomos derivados de um tipo peculiar de anticorpos apenas constituídos por cadeias pesadas, existentes em alguns membros da família *Camelidae*. Estes formatos naturalmente apresentam características biofísicas favoráveis e aptidão para reconhecerem epítopos não-convencionais. A quimera aqui desenvolvida foi capaz de entregar especificamente o efetor de RNA de interferência a células contendo o recetor CXCR4 na sua superfície. A funcionalidade da molécula de RNA entregue por endocitose foi também comprovada pelo silenciamento do promotor de HIV e da inibição da replicação viral em células portadoras do mesmo recetor. Adicionalmente, os resultados deste capítulo revelaram a aplicação deste "nanobody" numa estratégia de funcionalidade dependente da via de entrega. Nesta abordagem alternativa, um fator de transcrição constituído por uma proteína dedos-de-zinco foi conjugado com dois "nanobodies", um que reconhece o recetor CXCR4 e outro irrelevante. Alguns destes fatores de transcrição artificiais baseados na tecnologia de proteínas dedos-de-zinco possuem a capacidade inata de transpor a membrana celular e ganhar acesso ao meio intracelular, não necessitando de um método de translocação. No entanto, esta via de entrada é inespecífica, levando a um processo de distribuição ineficiente e em alguns casos indesejado. Os resultados deste estudo indicam que das duas quimeras apenas o fator de transcrição entregue por endocitose através do recetor CXCR4, em vez da via normal não específica de entrada destes fatores, funcionou como repressor da expressão do genoma do HIV a partir do promotor viral.

Em resumo, os resultados obtidos nesta dissertação de doutoramento demonstram o potencial dos anticorpos de domínio único como agentes de terapia da infeção por HIV, quer através da sua atividade neutralizante como inibidores de fusão ou na entrega de novas classes de moléculas terapêuticas. Adicionalmente, esta dissertação fornece informação relevante sobre novos alvos terapêuticos e a construção de potentes e promissores inibidores antirretrovirais.

Palavras-chave: Anticorpo de domínio único, HIV, Inibidor de fusão, Entrega específica, "Nanobody"

ABBREVIATIONS

Aa	Amino acid
AAV	Adeno-associated virus
AGO2	Argonaute 2
AIDS	Acquired Immunodeficiency Syndrome
APC	Allophycocyanin
ATF	Artificial transcription factor
aTTP	Acquired thrombotic thrombocytopenic purpura
BCR	B-cell receptor
bNAb	Broadly neutralizing antibody
BSA	Bovine serum albumin
CA	Capsid
CAR	Chimeric antigen receptor
cART	Combination antiretroviral therapy
CCR5	CC-chemokine receptor 5
CD4	Cluster of differentiation 4
CD4bs	CD4 binding-site
CDR	Complementary determining region
CDR1	Complementary determining region 1
CDR2	Complementary determining region 2
CDR3	Complementary determining region 3
CH	Heavy-chain constant domain
Chol	Cholesterol
CL	Light-chai constant domain
CRF	Circulating recombinant form
CRISPR/Cas9	Clustered regularly-interspaced short palindromic repeats/CRISPR-associated protein 9
CXCR4	CXC-chemokine receptor 4
CTD	C-terminal repeat domain
C3PO	Component 3 promoter of RISC
DART	Dual-affinity re-targeting
dsRNA	Double-stranded RNA
DVN	Degenerate codon
d1EGFP	Destabilized variant 1 of EGFP reporter
EC_{50}	50% effective concentration
E. coli	*Escherichia coli*
e-gp41	Gp41 ectodomain
Env	Envelope glycoprotein
ER	Endoplasmatic reticulum
ESCRT	Endosomal sorting complexes required for transport
Fab	Fragment antigen-binding
FBS	Fetal bovine serum
Fc	Fragment crystallizable
FDA	USA Food and Drug Administration
FITC	Fluorescein
FP	Fusion peptide
Gag	Group-specific antigen
H	Heavy-chain

HA	Hemagglutinin A
HCAb	Heavy-chain only antibody
HEL	Hen egg-white lysozyme
HIV	Human Immunodeficiency Virus
HIV-1	Human Immunodeficiency Virus type 1
$HIV\text{-}1_{NL4\text{-}3}$	HIV-1 laboratory-adapted strain NL4-3
HIV-2	Human Immunodeficiency Virus type 2
HR1	Heptad repeat 1
HR2	Heptad repeat 2
IC_{50}	50% inhibitory concentration
Ig	Immunoglobulin
IgA	Immunoglobin class A
IgD	Immunoglobin class D
IgE	Immunoglobin class E
IgG	Immunoglobin class G
IgM	Immunoglobin class M
IMAC	Immobilized metal ion affinity chromatography
IN	Integrase
INSTI	Integrase inhibitor
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IU	Infectious unit
K_a	Association rate
K_d	Dissociation rate
K_D	Affinity constant
KRAB	Krüppel-associated box
LB	Lysogeny broth
LTR	Long terminal repeat
LUV	Large unilamellar vesicles
MA	Matrix
mAb	Monoclonal antibody
MFI	Mean Fluorescence Intensity
miRNA	Micro RNA
MLV	Multilamellar vesicle
MPER	Membrane proximal external region
MVC	Maraviroc
MW	Molecular weight
n.a.	Not applicable
Nb	Nanobody
NC	Nucleocapsid
NF-kB	Factor nuclear *kappa B*
NLS	Nucleolar localization signals
NRTI	Nucleoside/nucleotide reverse transcriptase inhibitor
nNRTI	Non-nucleoside reverse transcriptase inhibitor
ORF	Open reading frame
PBS	tRNA primer-binding site
PBMCs	Peripheral blood mononuclear cells
PI	Protease inhibitor
Pol	Polymerase
Pol II	RNA polymerase II

POPC	1-palmitoyl-2-oleyl-sn-glycero-3-phosphocholine
Pre-miRNA	Precursor miRNA
Pri-miRNA	Primary miRNA
PR	Protease
PrEP	Pre-exposition prophylaxis
P-TEFb	Transcription elongation factor b
RISC	RNA-induced silencing complex
RNAi	RNA interference
RT	Reverse transcriptase
SB	Super broth
scFv	Single-chain variable fragment
SD	Standard deviation
sdAb	Single-domain antibody
SDF-1α	Stromal cell-derived factor 1-alpha
SEM	Standard error of the mean
SHIV	HIV/SIV chimera
shRNA	Short-hairpin RNA
siRNA	Small interfering RNA
SMIP	Small modular immunopharmaceutical
SNALP	Stable nucleic acid lipid particle
SP1	Specificity protein 1
SPR	Surface plasmon resonance
Tat	Trans-activator of transcription
TAR	Trans-activating response element
TBP	TATA binding protein
$TCID_{50}$	50% tissue culture infectious dose
UNAIDS	The Joint United Nations Programme on HIV/AIDS
VH	Heavy-chain variable domain
VL	Light-chain variable domain
VP16	Viral protein 16
VS	Virological synapse
ZF-TF	Zinc-finger transcription factor
ZFP	Zinc-finger protein
6HB	Six-helical bundle

Amino acids

A	Alanine	L	Leucine
R	Arginine	K	Lysine
N	Asparagine	M	Methionine
D	Aspartic acid	F	Phenylanine
C	Cysteine	P	Proline
Q	Glutamine	S	Serine
E	Glutamic acid	T	Threonine
G	Glycine	W	Tryptophan

TABLE OF CONTENTS

FIGURES

TABLES

CHAPTER I

INTRODUCTION

1.1Antibodies: The New Era of Biologics

Derived from the Immunoglobulin (Ig) family, antibody molecules are heterodimeric glycoproteins, produced by fully differentiated B-lymphocytes during the adaptive immune response. Antibodies target "non-self" molecules (antigens) such as pathogens, allergens, and toxins by interacting with an immunogenic sequence—epitope—within the foreign substance. Due to high affinity and specificity towards antigen and favourable characteristics such as solubility, stability and residual toxicity, therapeutic antibodies have driven the pharmaceutical industry with a pronounced market growth rate of biologics area through representatives in oncology, inflammatory/autoimmune and infectious disorders. In the last few years, antibodies volume approved and in late-stage development increased exponentially, representing ~75% of biopharmaceutical market in USA and Europe (BioPharma-Reporter.com). There are currently ~50 therapeutic antibodies for market sale, with seven ones being approved only in the former year, and over 280 in Phase II/III clinical studies.[1]

1.1.1 Antibody Structure and Function

In mammals' immune system, the antibody format presents two identical heavy chains (H) and light chains (L) in Y-shaped structure (Figure 1.1). The L chains belong to kappa (κ) or lambda (λ) subtypes and H chains to α, δ, ε, γ or μ isotypes. The isotype is determined by sequence and length of heavy-chain constant domains, distributing antibodies among five classes: IgA, IgD, IgE, IgG and IgM, respectively.[2] Because of its predominance in human serum, importance for immune response and excellent specificity characteristics, IgG constitutes the dominant format in immunotherapy and consequently the one being further addressed in the present thesis. In IgG class, each heavy chain is organized in three constant domains (CH1, CH2 and CH3) and one variable domain (VH), whereas the light chain is constituted only by a constant (CL) and a variable domains (VL). Variable domains are positioned into antibody N-terminus and together with the CL and CH1 regions constitute the fragment antigen-binding (Fab). This fragment is linked by a flexible sequence (hinge) to the CH2 and CH3 domains, constituents of the fragment crystallizable (Fc), where carbohydrates motifs are attached. Several inter-domain disulfide bonds along with the highly conserved intra-domain ones maintain antibody integrity.[3]

Variable domains mediate antibody specificity and affinity towards antigen mainly through three hypervariable loops, named complementary determining regions (CDRs; H1, H2, H3, L1, L2 and L3), with the six CDRs of each antibody "arm" setting up the antigen-binding site (Fig. 1.1). CDR loops are supported by four relatively conserved β-sheet strands, the framework sequences. The Fc region is the major responsible for antibody effector functions through complement and gamma Fc receptor (FcγR) binding. On the other hand, this IgG portion prolongs antibody half-life through a recycling mechanism dependent on neonatal Fc receptor binding.[4] Considering the scope of the present thesis, this introduction section will be focused on antibody variable domains.

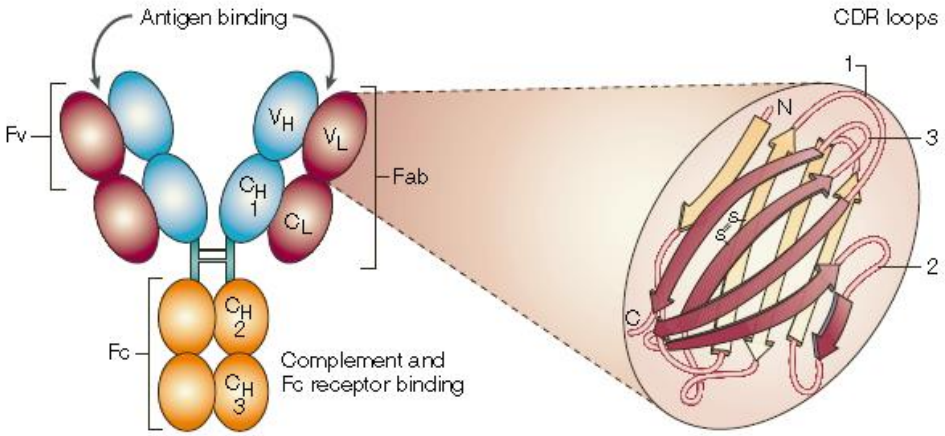


Figure 1.1: Antibody IgG structure.
Antibody molecule is divided into three major domains: two fragment antigen-binding (Fabs) and one fragment crystallizable (Fc). Each Fab comprises a variable fragment (Fv), constituted by antibody variable domains (VH and VL), one heavy-chain constant domain (CH1) and one constant light-chain variable domain (CL). Fc region comprises two heavy-chain constant domains (CH2 and CH3), represented in orange, responsible for complement and Fc receptor binding. The hinge region is represented in green. The inset represents a light-chain variable domain with the three complementary determining regions (CDR loops 1, 2 and 3). The arrows depict the β-sheet strands that form the CDR-supporting regions (frameworks). The conserved disulphide bond within each variable light-chain is also represented. Adapted from Brekke & Sandlie.[5]

1.1.2 Genetic Organization and Rearrangement of Antibodies

During B-cells development, several mechanisms ensure the generation of an enormous repertoire of antibodies with ability to target highly diverse antigens (for review see Schroeder & Cavacini).[2]

Antibody units (heavy-, κ light- and λ light-chains) are codified by multigene families divided into sets of tandemly arranged segments and situated on distinct chromosomal locations within the human genome. The diverse variable domains are assembled by combinatorial rearrangements of variable (V) and joining (J) genomic elements in case of light-chains and V, diversity (D) and J segments in case of heavy-

chains. This process, termed somatic recombination, constitutes the prime source of antibodies variability. Additional diversity in form of nucleotide insertions and deletions occurs at segment junctions during V(D)J joining (junctional diversification). After recombinant connection of heavy and light chains, another source of antibody diversification, the fully assembled immunoglobulins are expressed as B-cell receptors (BCRs) at surface of newly generated mature cells, constituting the naïve antibody repertoire.

In the second phase of B-cells development, lymphocytes harbouring antigen-specific immunoglobulins proliferate (clonal expansion) and undertake the second major mechanism of variability, somatic hypermutation, where randomly mutations occur at high rate in antibody variable domains. This antigen-driven activation of B-cells fine-tunes the antibody response. Clones presenting beneficial mutations towards high-affinity antigen interactions undergo preferentially expansion and survival (affinity maturation). Along with B-cells activation, class-switch recombination originates a pool of antibodies with distinct heavy-chain isotypes and consequently variable biological properties. Finally, B-lymphocytes differentiate into long-lasting memory cells, responsible for rapid recall of immunologic response against the same antigen, or antibody-producing plasma cells.

1.1.3 The *in vivo* Mechanisms of Action of Antibodies

Antibodies exert their function *in vivo* through three main mechanisms: neutralizing/antagonist activity; receptor signalizing/modulation and cellular depletion through antibody-dependent cellular cytotoxicity (ADCC) or phagocytosis (ADCP) and complement-dependent cytotoxicity (CDC) pathways.[6] In the former action mode, antibodies directly bind to ligands such as cytokines, growth factors, pathogenic agents, and toxins, preventing the interaction of these molecules with their cognate receptors. Antibodies can also function as antagonist effectors, blocking receptor activation/function. On the other hand, receptor-antibody interactions may also induce the downregulation of the former expression at cell-surface, diminishing available receptors for activation, or modulate cellular fates/functions through activation cascades. Internalization of cellular receptors can as well be an indirect consequence of antibody-mediated sequestration of respective ligands.[7] The depletion of antigen-bearing cells includes the Fc-mediated antibody properties. These functions rely on Ig cross-linking

with cell-surface receptors and FcγRs on immune effector cells such as natural killer (ADCC) and phagocytic (ADCP) cells or on complement components (CDC), leading to the destruction, phagocytosis or lysis of target cells, respectively.

1.1.4 Antibody Engineering: The Rise of Single-domain Antibodies

In contrast to diversity of *in vivo* antibody response, specificity towards a single epitope is critical for safety and efficacy in therapeutics. Hybridoma technology enabled the production of the first pools of mouse antibody clones (monoclonal antibodies; mAbs) through fusion of antibody-producing cells and an immortalized cell line.[8] Afterwards, DNA recombinant technology enabled the development of chimeric and humanized mAbs, which circumvented lack of effector functions as well as reduced immunogenic elicitation[9] and half-life issues of mouse mAbs when applied in human therapeutics.[10,11] Finally, advanced strategies of genetic engineering including selection of antibodies from phage-displayed human libraries[12] or immunized transgenic animals[13] and human memory B-cells immortalization enabled the production of fully human mAbs, overcoming the remaining immunogenicity of chimeric and humanized mAbs (Fig. 1.2).

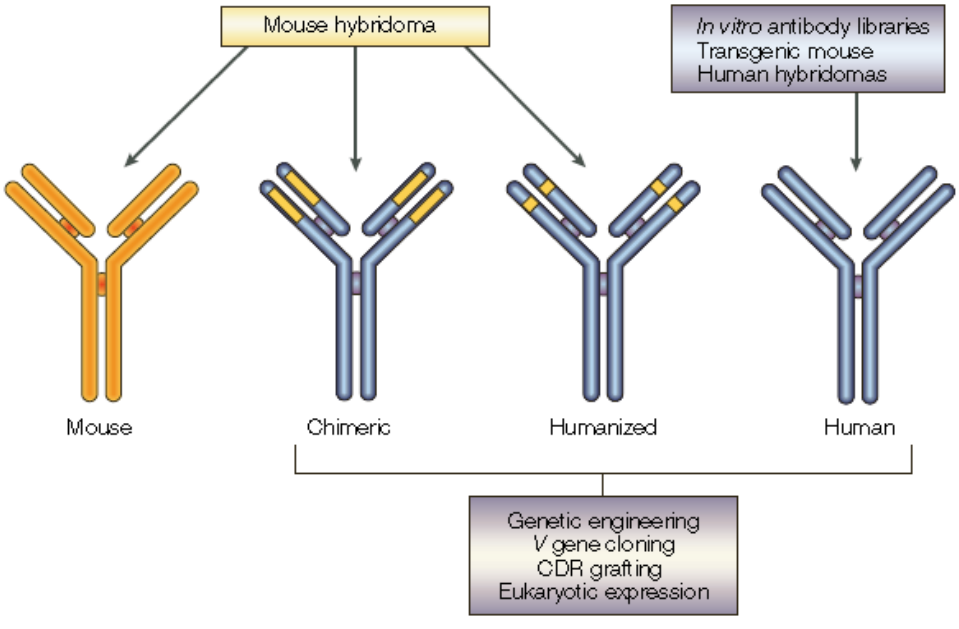


Figure 1.2: Evolution of antibody engineering.
Yellow colored antibody represents mouse-origin, whereas blue color stands for human-origin. Mouse mAbs were produced by mouse hybridoma, being further genetic engineered through V gene cloning or CDR grafting into antibodies with mouse-origin restricted to variable regions or CDRs, chimeric and humanized mAbs respectively. Chimeric or humanized antibodies were generated and expressed in eukaryotic cells. Fully human antibodies are selected by screening of in vitro antibody libraries or generated by hybridomas, where B-cells isolated from immunized transgenic mice for the human Ig locus are fused with myeloma fusion partners. Adapted from Brekke & Sandlie.[5]

Despite the quite successful application in therapeutics, large molecular weight (MW) of conventional IgG format impairs the targeting of sterically occluded epitopes and penetration into densely packed tissues. To overcome these major issues, smaller antibody scaffolds were developed such as Fab, single-chain variable fragment (scFv) and single-domain antibody (sdAb) (Fig. 1.3). Fab (~50 kDa) had previously been defined has containing the variable and CH1 domains of an antibody, whereas scFv (~30 kDa) is constituted by a VL-VH pair connected by a flexible linker, commonly $(GGGGS)_x$.[14] In case polypeptide linker is absent or reduced, antibody fragments are denominated Fvs and tend to multimerise into diabodies and triabodies.[14] sdAb constitutes the smallest functional fragment (~15 kDa), being constituted by the antibody VL or VH domains, and the only one whose antigen-binding site is restricted to three CDRs. Nevertheless, sdAbs demonstrated an affinity and specificity similar to the whole antibody format.[15] Moreover, fully-functional antibodies devoid of light chains were identified as natural constituents of camelid and shark immune systems.[16]

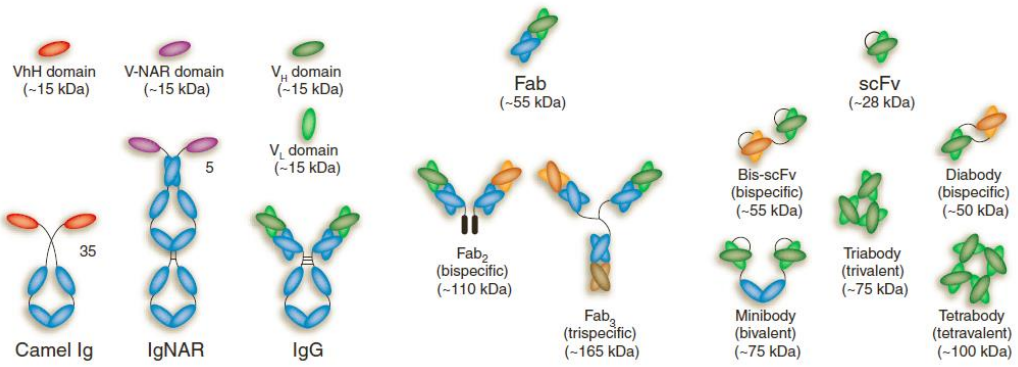


Figure 1.3: Antibody fragmentation and alternative engineered variants.
Left) Standard IgG format compared with camel IgG and shark IgNAR counterparts and their respective V domains. **Right)** Most common antibody fragments and their respective molecular weights as well as multimerize versions of the same. Adapted from Holliger & Hudson.[17]

Innovative scaffolds of miniaturized antibodies have also been engineered by assembling distinct IgG formats and/or portions in one single molecule. Some examples are IgG-scFv fusion, scFab (single chain Fab), $[scFv]_2Fab$, scFv-Fc, $[(scFv)_2Fc]_2$ minibody/small modular immunopharmaceutical and di-dibodies (Fig. 1.4).

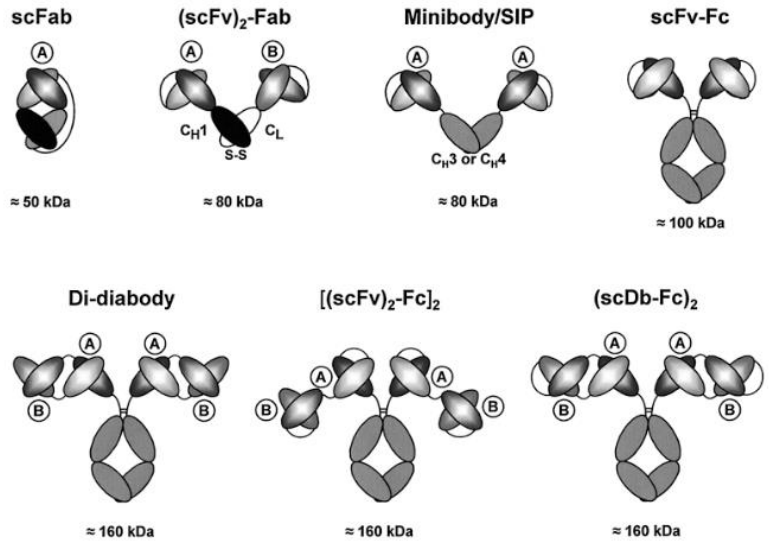


Figure 1.4: Alternative versions of antibody fragments.
Adapted from Little.[18]

1.1.4.1 Single-domain Antibodies as Therapeutic Molecules

Single-domain antibodies, similarly to other antibody fragments, exhibit remarkable therapeutic advances relative to the classical IgG structure, in addition to the well-known biophysical properties shared with full-size antibodies. Minimal size of domain antibodies enables access to novel and cryptic epitopes, as surface proteins of infectious virus, and enhances tissue penetration/distribution, for example in tumours and blood-barrier crossing. Moreover, structural simplicity of single-domain antibodies allows a cost-effective manufacturing on prokaryotic systems as well as affinity and avidity improvements by basic engineering techniques. For example, individual sdAbs can be easily assembled in multimers (dimers, trimers and tetramers), where specificity could differ within each module.[19] This formatting also contributes to half-life extension of sdAbs—rapid cleared from circulation because of low renal threshold. Nonetheless, other strategies can be applied to improve pharmacokinetic profile of sdAbs, including conjugation or fusion to albumin (plasma protein with extraordinary extended circulation time), polymers, carbohydrates or albumin-binding motifs.[20] Domain antibodies can as well be re-formatted as Fc-bearing molecules, in cases Ig effector functions are desirable or even mandatory, although advantage conferred by microorganism production be compromised. Alternatively, sdAbs can be combined with small active moieties. Moreover, these effector domains could contribute to novel and specific/potent actions that surpass traditional antibody functions provided by the Fc portion. Active moieties include radionuclides, cytokines, toxins, enzymes, peptides, and proteins. In case effector

functions are dispensable, the absence of Fc receptor as occurs in minimal antibody fragments could provide an additional safety benefit through avoidance of immunogenicity, a central issue on full-IgG therapeutic applications.

Given the relative conservation of framework regions, binding characteristics of the variable domains rely almost exclusively on CDRs sequence and length, being the second factor a primary determinant of binding-site shape. Nonetheless, majority of natural variable domains present length-constraint CDRs, only adopting a limited set of canonical conformations and consequently presenting a restricted capacity of antigen recognition.[21] Shorter loops favour the recognition of small antigens as haptens and peptides through deep pockets and grooves topographies (convex binding-sites), which contrasts with the more flat and protruding paratopes (longer loops) that preferentially target large antigens as proteins. Binding-site length also differ within each loop and chain type with L1, L3, H2 and H3 CDRs assuming an additional role on paratope topology. CDRs of light-chain domains commonly range from 7-13 aa, with some VLs presenting until 17/18 residues in L1 and/or L3. Binding-sites from heavy-chains mainly differ from 5-19 aa with some VHs presenting until 38 aa.[21,22] Amino acid distribution also dictates preferential binding of CDRs and, similarly to length, each loop has its own unique composition and contact preferences.[22,23] In general, tyrosine and tryptophan are the most common residues found on antibody binding-sites followed by isoleucine and serine. The remaining representativeness are more correlated with antigen type. For example, glycine, valine, cysteine, lysine, glutamate and phenylalanine residues are generally under-represented in protein and virus binders, whereas leucine and methionine are usually over-presented. In contrast with histidine, leucine and asparagine, alanine residue is most rarely found on haptens binding-sites. Residues position on binding-sites is also a factor on Ag-Ab interactions with the central amino acids, which are also the most variable ones, being the most involved in epitope interactions. The non-contacting residues are determinants for maintaining the conformational configurations of loops. Despite only 3% of sequenced variable domains being from non-human and -mouse origin, CDRs from distinct species also differ in binding and physical-chemical properties. For example, rabbit variable domains present greater variance in length and sequence as well as prevalence of long CDRs and a greater contribution of light chains for antigen binding than human and mouse antibodies.[24]

Antigen-binding strictly mediated by three CDRs such as in human-origin sdAbs exists naturally in antibodies belonging to *Camelidae* family and Cartilaginous fish class.[25,26] The variable domains of these heavy-chain only antibodies (HCAbs) are denominated VHH for camelids and VNAR for sharks (Fig. 1.5). In contrast to the human gamma heavy chain disorder, where genetic deletions in VH and CH1 domains result in non-functional HCAbs,[27,28] these unique antibody formats are capable of antigen recognition and binding. Moreover, extensive studies revealed peculiar chemical and biophysical characteristics of camelid-based VHs. The extension and structural flexibility of camelid CDRs compensate variability restriction from the absence of VL domains and suit propensity for cleft binding, a highly desirable feature to target structurally restricted regions oftentimes related with immune-evasion epitopes.[29–31] On the other hand, these CDR features increase array conformations of VH domains, namely to non-canonical architectures, often improving VHH-antigen fit and consequently binding affinity.[30] VHHs are as well stable and soluble domains presenting no signs of aggregation when produced individually (nanobodies), which contrasts with the early isolated VH human domains.[15] A major contributor to nanobody stability is the dissimulation of VH-VL interface through the presence of hydrophilic amino acids in conserved positions of this region.[32] Extended CD3 can also bend backwards further shielding VL/VH region by interactions with framework residues[30] although the expected negative impact of this feature in CDR3 diversity.[33] An additional disulfide bond connecting CDR3 and CDR1 or CDR2 loops, common in VHH domains, contributes as well to the stability of these non-standard VHs. Shark and camelid-based VHs share characteristics in terms of structure, biophysical properties and CDRs extension and plasticity.[34–37] However, camelid domains have been further studied and therapeutically evaluated because of a higher homology degree with human VHs and easiness in isolation from naïve and immune libraries sources.

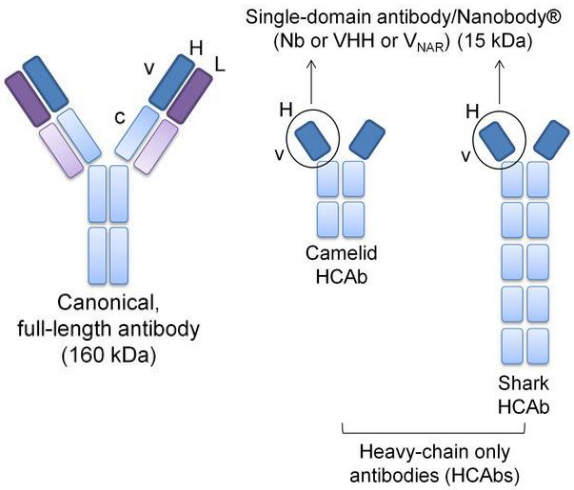


Figure 1.5: Structural comparison of canonical antibody format with camelid and shark heavy-chain only antibodies (HCAbs).
Heavy-chains (H) are represented in blue and light-chains (L) in purple, with the constant (c) domains in lighter color and variable (v) ones in a darker version. In addition to light-chains, camelid HCAbs lack CH1 units. On the other hand, heavy-chains of shark HCAbs present five constant domains. Adapted from Doshi et al.[38]

Despite tremendous stability and solubility along with remarkable refolding properties after exposure to high temperature, low pH stability and enzymatic degradation, VHHs are prone to immunogenicity when applied as therapeutic agents although no immune response was raised so far when these non-human VHs were applied in mouse and human therapy. In consequence, "camelization" strategies have been applied to engineer human VH domains[39–44] in an attempt to minimize aggregation propensity[15,45–48] and mimic biophysical properties of VHHs in terms of solubility, heat refoldability and purification yield.[45] These strategies mainly consist in mutations to increase hydrophilicity of VL/VH interface and addition of non-native disulfide bonds.[49–54] Mutational changes on framework residues outside VL/VH interface or even CDRs itself constitute alternative approaches of "camelization",[33,55] sometimes resulting in engineered domains that rival the stability and cleft recognition of camelid ones.[56] In alternative, stable human VH domains have been directly selected from antibody-based libraries, mainly through phage display technology.[57–62]

Only sdAbs in form of VHH or human VH domains are currently under clinical trials development. A diverse array of camelid-based VHs are being studied against distinct targets within autoimmune, inflammatory, cardiovascular, respiratory, oncologic and neurologic diseases. Indicated for treatment of acquired thrombotic thrombocytopenic purpura (aTTP), caplacizumab is the leading VHH in product development pipeline (Table 1). A human domain antibody (GSK2862277) against tumor necrosis factor

receptor 1 is under Phase II evaluation for preventing acute lung injury and an albumin-binding domain antibody (NCT02829307) has also been evaluated in Phase I as part of a strategy for half-life extension of rapid cleared drugs.

Table 1.1: VHH domains under clinical development as December 2016.
Adapted from Ablynx Inc.

Although VH domains have been widely explored for human therapeutics, some studies demonstrated that VLs can present high expression yield, non-aggregation propensity[46–48,63–68], reversibility of thermal unfolding and resistance to gastrointestinal proteases, in some cases better than VH domains.[69] Previously, Brinkmann *et al.* had already described a VL domain as more active and less prone to aggregation than its VHs counterpart.[70] These findings highlight the gap in exploration of VL formats potential for therapeutic applications.

1.1.4.2 Phage Display Technology and Synthetic Libraries Tailoring

Antibody development largely depends on implementation of display platforms. These *in vitro* technologies for the isolation of therapeutic candidates consist in three main steps: generation of a large repertoire of individual clones (library); display and selection rounds towards the target or desired property and functional screening and characterization of selected compounds (Fig. 1.6). Major display technologies include ribosome and mRNA/cDNA display, phage display and cell display.[71]

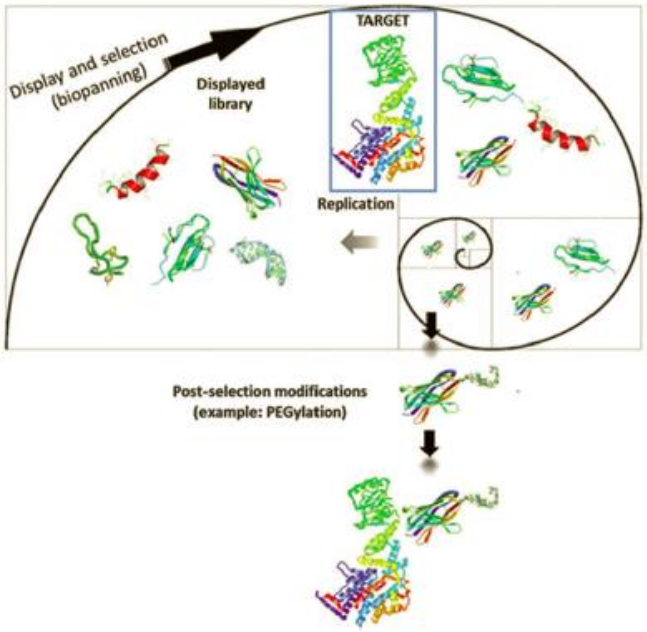


Figure 1.6: Display technologies flow-through.
Iterative display and selection cycles (biopannings) diminish displayed-library size, being enriched for binders of interest targets. Post-selection modifications can complement function of selected binders or improve their biophysical properties. Adapted from Gálan et al.[71]

Phage display is the most well-established technology for *in vitro* selection of antibody formats, relying in the presentation of functional proteins in fusion with viral coat proteins at surface of filamentous bacteriophages (phages).[72] Genes coding for the surface-exposed molecules are encapsulated within the phage particle, establishing the phenotype-genotype linkage essential for the success of display technologies. In each selection cycle, phage-displayed library is subjected to selective pressure against a target molecule, usually immobilized onto a solid surface. Non-specific phages are washed out under stringency washing conditions, while phages of interest are eluted, enriched by infection and re-growth in bacteria (Fig. 1.7).

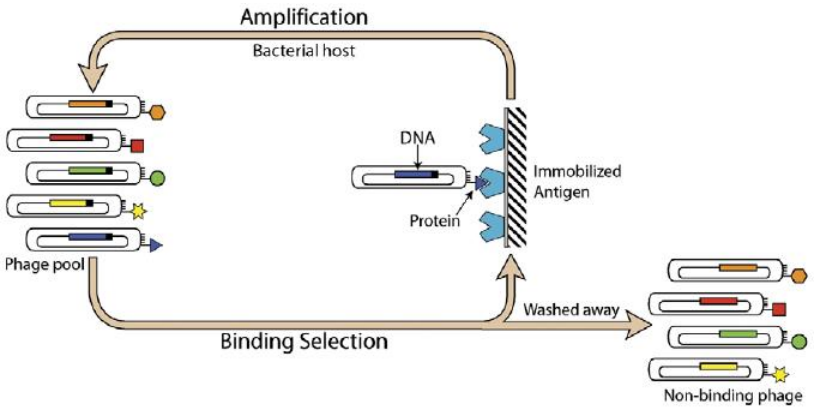


Figure 1.7: Flow-through of phage-display selection cycle.
A phage pool with great diversity containing the surface-displayed molecules and respective DNA coding sequences within the phage particles are incubated with an immobilized antigen (binding selection). Non-binding phages are

washed way, while specific ones remain bound to the antigen. These phages are afterwards eluted, being amplified in a bacterial host. Adapted from Sidhu & Koide.[73]

Diverse types of antibody libraries can be subjected to *in vitro* selection and screening. A naïve library represents the natural immune repertoire of a donor, which variability relies only in sources of sequence variation during B cells maturation, mainly the rearrangements of V(D)J genes. Secondary lymphoid tissues harbor antigen-activated B cells, being a preferential source for additional diversity in a naïve repertoire. These libraries enable selection and isolation without knowledge of antigen structure or sequence. On the other hand, they present restrictions in selection of antibodies against non-canonical targets and heterogeneous representativeness of some gene combinations.[74–76] For immunized libraries, recovery of plasma and activated B-cells occurs 5-6 weeks after animal inoculation with the desirable antigen. These libraries enable to direct the selection process towards current needs, although requiring antigen knowledge. Nonetheless, immunized repertoires limit selection of antibody clones targeting toxic, non-immunogenic, and self antigens in addition to non-standard ones. Immunization is as well a time-consuming process and results in an unpredictable antigen response. Due to ethical concerns, B cells collection is restricted to patients and non-human animals. To overcome the limitations of naïve and immunized repertoires in antigen targeting, CDR sequences can be designed *in silico* and cloned into known antibody scaffolds, originating synthetic and semi-synthetic libraries. Synthetic libraries enable precise control over library variability and representativeness, providing *de novo* diversity and conformations, in addition to capability for *in vitro* screening of all antigen types.[77] Nonetheless, the synthetic library should be rationally designed to maximize functional diversity and consequently library quality. This process is expected to reduce representativeness of misfolded, unstable, nonproper displayed and nonfunctional clones. Several strategies are available to the design/randomization of synthetic CDRs (reviewed in Neylon).[78] Briefly, these methods introduce diversity in form of point mutations and insertions/deletions that can occur randomly or in specific positions of a target gene. In alternative, randomizing methods do not directly create *de novo* diversity, but combine the existing one in novel variants. At Fig. 1.8 is represented one technique from each category of CDRs design.

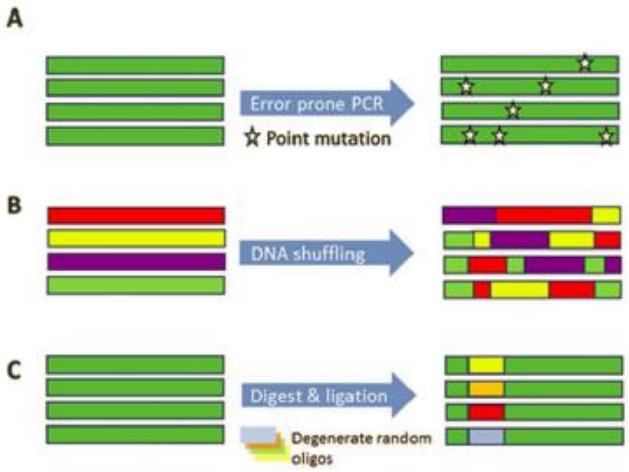


Figure 1.8: Strategies to design of synthetic CDRs.
A) Error prone PCR introduces point mutations randomly within the target DNA being copied. **B)** DNA shuffling technique mixes fragments of distinct DNA sequences in novel combinations. **C)** Digestion & ligation technique introduces sequences of degenerate codons at specific positions of gene sequence. Adapted from Gálan et al.[71]

CDRs tailoring for construction of synthetic libraries not only allows full control over clonal diversity as also gives insights into depth-knowledge of the molecular recognition/binding properties of natural binding-sites. For example, high-affinity clones with distinct specificities were selected from libraries with serious chemical restricted diversity.[79–82] Consequently, it was concluded that only few amino acids are main contributors for CDR-mediated binding in minimalist synthetic proteins.[83] One of these residues is tyrosine, whose structure in aromatic ring dominates the antigen interactions establishment. Previously, certain residues such as serine, glycine, histidine, arginine and tryptophan had already been described as substantially more frequent at binding sites than others.[84,85] Overall, these studies outstand the targeting diversity of distinct antibody formats by using a limited set of residues.

1.2 Human Immunodeficiency Virus

Despite remarkable therapeutic advances, Acquired Immunodeficiency Syndrome (AIDS) continues to be a major public health and economic burden. According to UNAIDS (The Joint United Nations Programme on HIV/AIDS), ~36.7 million people were infected with HIV at the end of 2015 and ~1.1 million people died from AIDS only in this year. AIDS is characterized by an insidious deterioration of the cellular immune system, where a progressive depletion of $CD4^+$ T-cells ultimately leads to the appearance of opportunistic infections (AIDS phase). Human Immunodeficiency Virus (HIV) is the etiological agent of AIDS,[86,87] being divided in two types: HIV-1 and HIV-2.[86–88] These

viruses differ in disease severity, evolution and transmissibility as well as prevalence and distribution worldwide. HIV-1 causes a more severe syndrome with a faster evolution towards the fatal AIDS phase and presents a higher transmission rate and prevalence. In contrast to worldwide distribution of HIV-1, HIV-2 is almost confined to West Africa even though the number of cases had lately increased in Europe, India and USA.[89] HIV-1 is classified in four groups: Major or Main (M), non-M and non-O (N)[90], Outlier (O) and P[91]. The M group is the most prevalent and can be subdivided in clades from A-D, F-H, J and K.[92] These genetically distinct subtypes can in turn recombine and originate circulating recombinant forms (CRFs). HIV-2 is distributed through groups A to H, but only A and B cause AIDS epidemics, being the A group the far most prevalent worldwide. CRFs were also identified for HIV-2.[93–95] HIV presents high variability intra and inter-clades, a major reason for challenge design of drugs and vaccines, with an average nucleotide diversity of genomes being 50% between HIV types, 37.5% among HIV-1 groups, 14.7% among clades, 8.2% within individual clades and less than 1% within single patients.[96]

1.2.1 HIV morphology and genetic structure

HIV-1 belongs to the Lentivirus genus of the *Retroviridae* family with a diploid genome of two positive single-strand RNA molecules that are reversely transcribed to a double-strand DNA molecule and integrated into host cell genome. HIV-1 also presents a lipid envelope acquired from the host cellular membrane during the virus budding. In this structure are embedded the viral envelope glycoprotein (Env) and several host-derived membrane proteins.[97] HIV-1 genome (Figure 1.9) codifies nine open reading frames (ORFs) flanked by long terminal repeats (LTRs) with the 5'LTR containing the HIV promoter.[98–100] The *group-specific antigen* (*gag*), *polymerase* (*pol*) and *env* ORFs are translated as polyproteins precursors, with structural proteins being only codified in the *gag* and *env* genes. The *gag* gene originates the matrix (MA), capsid (CA), nucleocapsid (NC), p1, p2 and p6 proteins. Matrix protein overlays the lipid envelop internally, while capsid proteins shields the viral core through a coned-shaped structure and nucleocapsid units coat the HIV genome. ORF *env* encodes the gp160 precursor of Env, cleaved and processed during exportation towards cell-surface to originate gp120 and gp41 subunits. Finally, the *pol* gene codifies the reverse transcriptase (RT), integrase (IN), and protease (PR) enzymatic proteins. The remaining six ORFs of the HIV-1

genome encode for viral accessory (Nef, Vif and Vpr) and regulatory (Vpu, Tat and Rev) proteins, being the last three not incorporated into HIV virion. In case of HIV-2, Vpu is substituted by Vpx (Fig. 1.9).

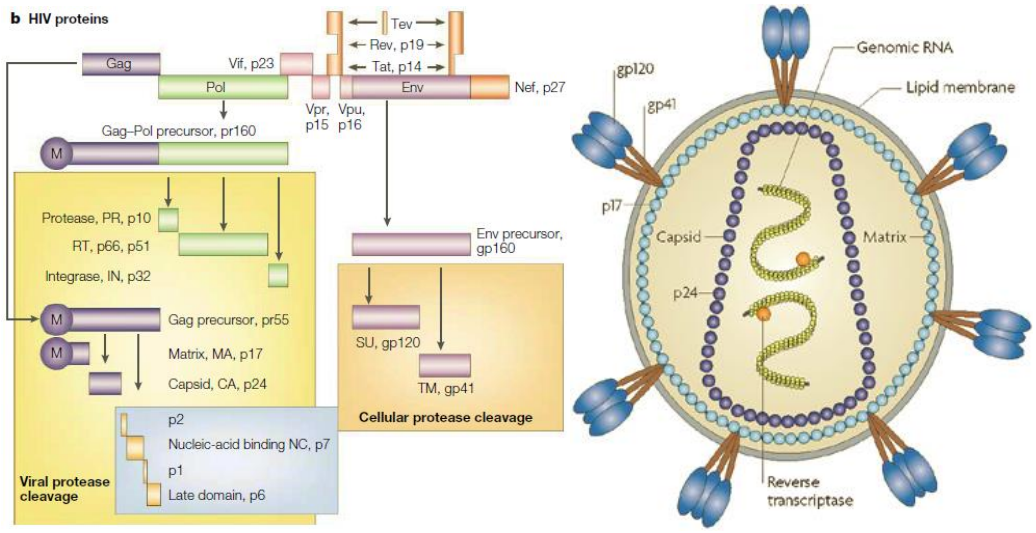


Figure 1.9: Structure of HIV-1 genome (A) and organization of HIV-1 virion (B).
Adapted from Shum et al. and Hedestam et al.[101,102]

1.2.2 HIV replication cycle

The replication cycle of HIV-1 is divided into infection and expression phases. Infection phase initiates with virus-cell interaction and concludes with the integration of proviral DNA into host genome. Expression of viral proteins triggers the second phase, which involves all steps until the budding and maturation of the virion (Fig. 1.10).

HIV replication cycle initiates with the fusion of viral envelope and host cell membrane by engagement of viral glycoprotein Env to cell-surface receptors. This process culminates in the liberation of viral core into cell cytoplasm. Subsequently to uncoating, RT enzyme converts HIV genomic RNA into a linear double-stranded DNA molecule, which is imported to the cell nucleus and integrated into host cell genome through viral integrase activity. Short mRNAs originate the early viral proteins such as Tat, which in turn are responsible for driving the expression of late/structural proteins from the integrated HIV genome (provirus) using the cellular machinery. After expression, all viral components must travel along cytoplasm to the assemble location in specialized rafts domains of cell membrane. Assembly and budding of the new-born virions characterize the last stage of HIV replication cycle. During or shortly after cell release, protease enzyme promotes the cleavage of viral precursors into the individual constituents of HIV virion. This final step converts immature virions into their mature

infectious form. Viral accessory and regulatory proteins such as Vif and Vpx/Vpu assures replication cycle completion by counteracting the activity of several host restriction factors in arresting HIV infection.

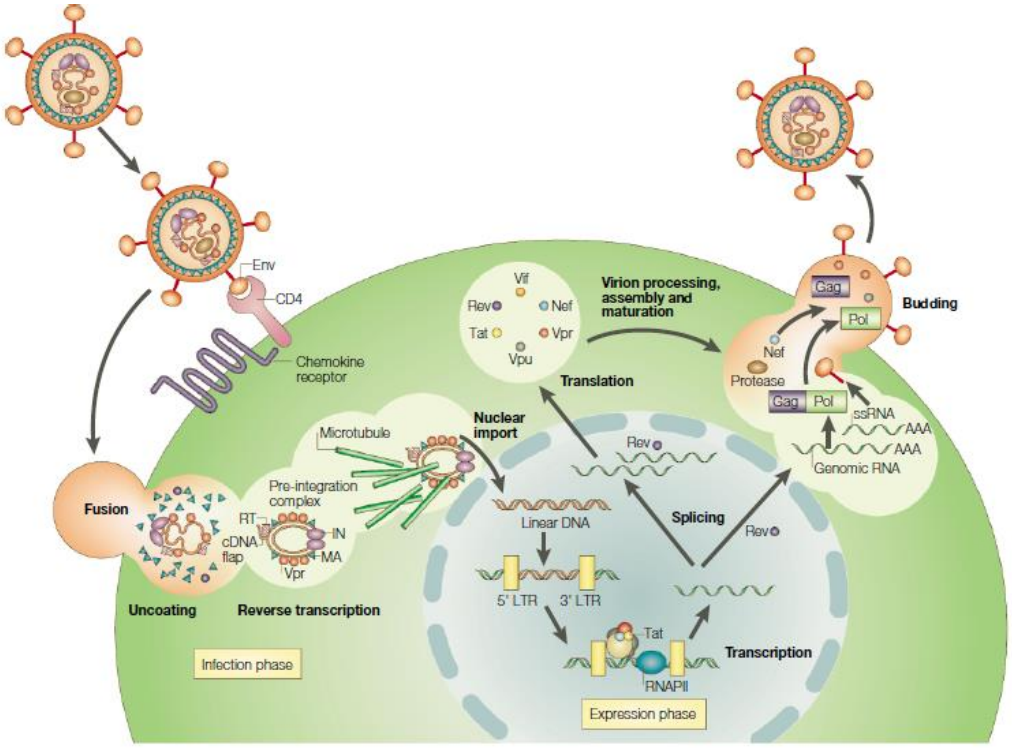


Figure 1.10: HIV-1 replication cycle.
Interaction of viral envelope glycoprotein (Env) with cellular CD4 receptor and a chemokine co-receptor results in the fusion of virion envelope and cellular membrane. Following the fusion step, viral core release in cell cytoplasm (uncoating) results in reverse transcription of HIV genomic RNA into a dsDNA molecule. This viral cDNA together with several viral and cellular proteins is assembled into a pre-integration complex, whose nuclear import is mainly mediated by Vpr viral protein. In the nucleus, viral DNA is inserted into host cell genome through integrase activity, which concludes the infection/early phase of HIV replication cycle. Together with viral protein Tat, the long terminal repeat (5' and 3' LTRs) flanks of provirus have critical roles in transcription of HIV mRNA. This transcript can codify the genome of new-born virions or suffer alternative splicing to originate all viral proteins and polyproteins precursors, after being nucleus exported through a Rev-dependent pathway. Processing, assembling and maturation of novel viral particles occurs at cell membrane mainly through Gag and Gag-pol polyproteins and protease viral protein. Adapted from Peterlin and Trono.[108]

The 5'LTR, where HIV promoter is allocated, is characterized by 3 regions: U3, R and U5. The U3 region contains the cis-acting regulatory elements of HIV promoter, being further divided into negative regulatory region, enhancer, and core promoter. Viral transcription initiates at R region and terminates with polyadenylation after R region of 3'LTR. Several cellular transcription factors and viral regulatory proteins that interact with U3 region of LTR regulating HIV transcription are represented in Fig. 1.11.

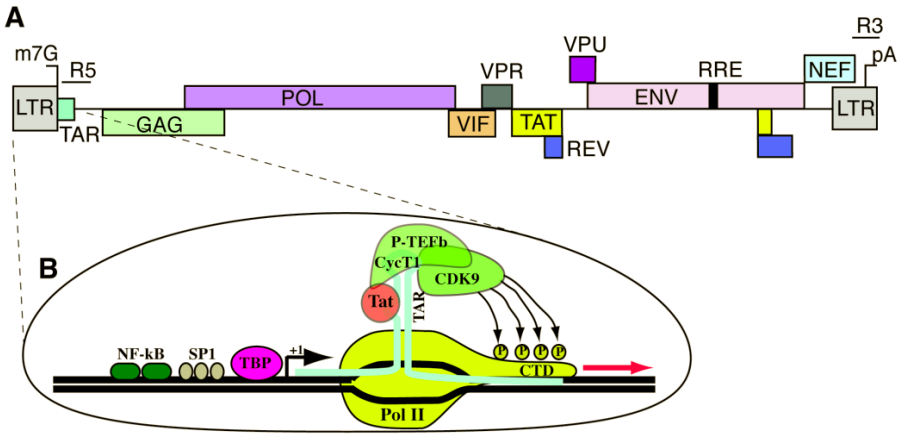


Figure 1.11: LTR transcriptional activation.
A) Localization of viral promoter on HIV genomic organization. **B)** The inset represents the viral transcription with cellular transcription factors such as NF-kB, SP1 and TBP and viral Tat binding to TAR sequence leading to P-TEFb assembly and consequent phosphorylation of RNA polymerase II (Pol II) CTD. Adapted from Caputi.[104]

1.2.3 HIV-1 entry

1.2.3.1 Env structure

HIV-1 envelope incorporates the Env glycoprotein responsible for HIV tropism and entry. Env is composed by a trimer of gp120 surface glycoprotein non-covalently associated with gp41 transmembrane protein.

Gp120 consists of five conserved (C1-C5) and five variable regions (V1-V5) distributed through two domains (Fig. 1.12). The conserved regions are concentrated in the inner domain, which constitutes the sterically occluded core of gp120, while variable regions—major contributors to immune-evasion—are highly glycosylated and mainly located at outer domain. Gp120 presents two crucial regions for HIV entry process, the receptor and co-receptor binding sites. CD4 membrane-spanning protein constitutes the prime HIV receptor.[105–108] CD4 binding-site (CD4bs) is constituted by a hydrophobic cavity between inner and outer domains of gp120, well-conserved and devoid of glycosylation, thus a favorable target for HIV inhibition. CD4-bearing cells are the primary target cells of HIV infection, preferentially T-lymphocytes, but also macrophages, monocytes, dendritic and microglia cells. Of note, CD4-independent HIV-1 infectivity has also been described.[109] Co-receptor binding site is located within V3 loop of bridging sheet domain, only set-up upon CD4-gp120 engagement. This fact constitutes an additional immune-evasion strategy of HIV. The major co-receptors of HIV are CXC-chemokine receptor 4 (CXCR4) and CC-chemokine receptor 5 (CCR5).

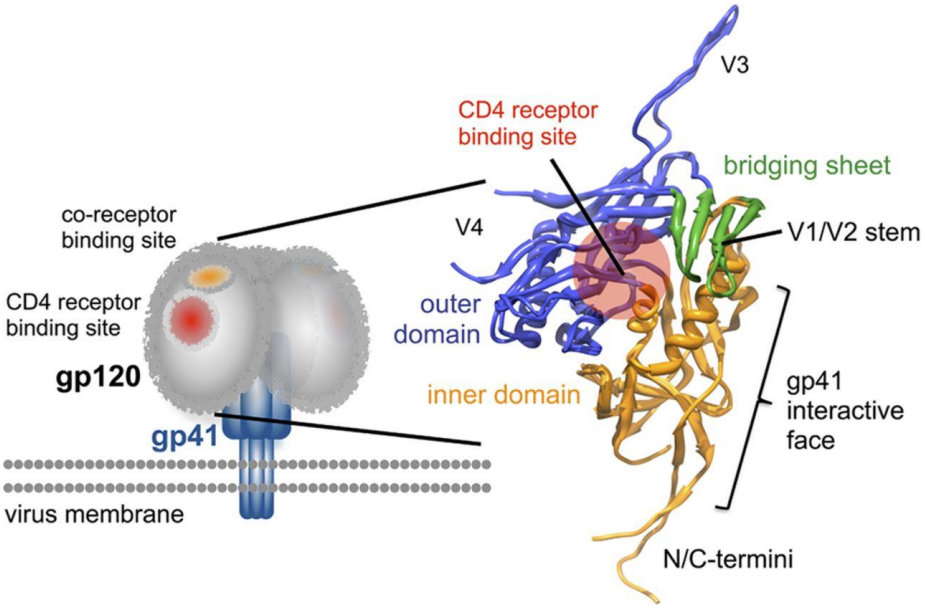


Figure 1.12: Structure of gp120 in CD4-bound conformation.
The blue region represents the outer domain, whereas the orange-colored chain represents the inner domain. The green region is the bridging sheet domain composed of two pairs of anti-parallel β-sheet strands (two from each domain) and V1/V2 loop from the inner domain, only set-up upon CD4-gp120 engagement. CD4 binding-site, depicted as a red circle, is located between the inner and outer domains. Gp120 interacts with gp41 mainly through the inner domain and N/C terminal extensions.[110,111] Adapted from Guttman et al.[110]

Gp41 is constituted by three major domains: cytoplasmic tail, transmembrane anchor and ectodomain (e-gp41). Extracellular portion is the main mediator of HIV entry process in gp41 context. Gp41 ectodomain is characterized by a hydrophobic N-terminal fusion peptide (FP), two α-helix heptad repeat regions 1 and 2 (HR1 and HR2) and a membrane proximal external region (MPER).

1.2.3.2 HIV-1 entry process: gp41 as fusion key

HIV-1 entry, the first phase of viral replication cycle, begins with adhesion of viral particle to host cell through interaction of membrane proteins incorporated into viral envelope and attachment factors in host cell-surface. These contacts promote spatial approximation between viral glycoprotein Env and HIV receptor and co-receptor at cell-surface. In a second phase, gp120 triggers the fusion process through binding to the cellular receptor CD4.[105–108] CD4 interaction leads to rearrangements in gp120 such as the formation of a bridging sheet domain and exposure of V3 loop, both important for cellular co-receptor binding.[112] Depending on the co-receptor used, HIV-1 is classified into R5 (CCR5 usage), X4 (CXCR4 usage) or R5X4 (both co-receptors usage) strains.[113] Alternative chemokine receptors were described as HIV-1 co-receptors, however there is no evidence for their role *in vivo*.[114,115] Co-receptor engagement results in gp120 dissociation from Env complex, leading to a cascade of conformational changes within e-

gp41. Fusion peptide is released and inserted into the host cell membrane, originating an unstable e-gp41 conformation, named fusion intermediate structure. In this transient structure that quickly collapses into a less-energy state, termed six-helical bundle (6HB), HR1 and HR2 regions are conformationally exposed. 6HB conformation is characterized by the position of the three α-helix HR1 regions into a parallel fashion, forming the conformation core, and the three HR2 regions packed in an oblique and antiparallel manner into the highly conserved and hydrophobic grooves of the HR1 peptides.[116–120] Transition of e-gp41 to 6HB conformation drives the viral fusion process by bringing the virus envelope and cellular membrane in apposition, mediated by the retraction of the HR1/HR2 trimer of heterodimers.[121] These alterations together with deformations in the plasmatic membrane triggered by gp41 FP and destabilization of the virus envelope by gp41 MPER lead to the formation of the fusion pore. The number of fusion pores needed to the successfully deliver of virus payload into the cells cytoplasm remains unclear. The HIV-1 entry process is represented in Fig. 1.13.

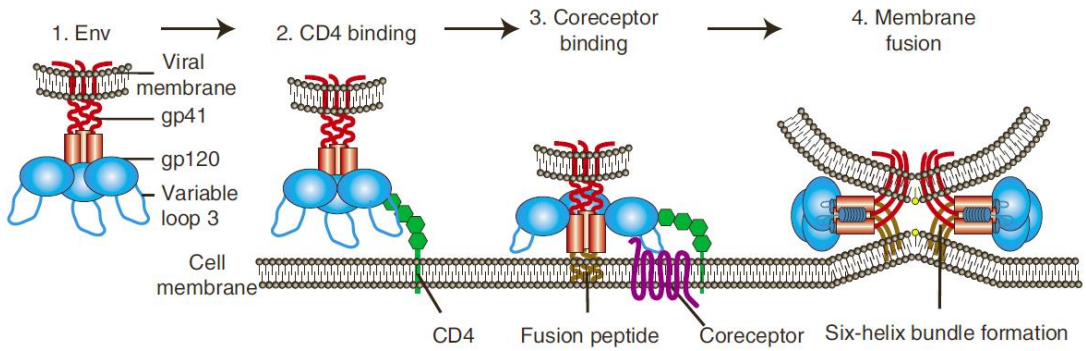


Figure 1.13: HIV-1 entry.
Viral Envelope (Env) is comprised by three gp120 and gp41 subunits **(1)**. Gp120-CD4 interaction triggers the fusion process **(2)**, causing conformational rearrangements in viral glycoprotein followed by co-receptor binding in variable loop 3 (V3) and insertion of gp41 fusion peptide into host cell membrane **(3)**. The six-helix bundle formation concludes the fusion process **(4)**. Adapted from Wilen et al.[122]

1.2.3.2.1 Cell-Cell Spread of HIV

HIV-1 transmission occurs through two routes: virus-cell and cell-cell. The virus-cell was extensively explained previously and it is the most well-known process. Cell-cell fusion is a more recently described HIV spread and its role and importance in the natural infection is still under debate. However, there are some studies that already demonstrated the efficiency and rapidity of this transmission mode *in vitro*, which anticipates a relevant if not dominant mode of the virus dissemination in infected individuals.[123–129]

Cell-cell fusion enables direct HIV spread from infected to non-infected target cells, establishing foci of infection. This route mainly occurs by formation of virological synapses (VS)[130–135] with alternative mechanisms such as transient cell-cell contacts and longer-range intercellular interactions including nanotubes and filopodia[136,137] contributing residually. In VS mechanism, cytoskeleton polarization towards cell junctions is triggered by gp120-CD4 interaction (donor-target cells) leading to the formation of specialized structures, where viral proteins are redirected and concentrated.[138] One question that remains unclear is related with HIV mechanism to penetrate target uninfected cell. Since HIV entry is pH-independent, this process could simply occur by fusion between adjacent target cells ("fusion-from-without"). However, evidences suggesting an alternative, or even required, endocytic pathway for HIV entry have been reported.[139–142] In this theory, virus fusion does not proceed beyond lipid-mixing step at cell membrane level, requiring the formation of an endosomal compartment for viral capsid release on cell cytoplasm.[140]

1.3 HIV/AIDS therapy

Combination antiretroviral therapy (cART) constitutes the approved scheme for HIV treatment on seropositive/AIDS individuals, being composed by several antiretroviral drugs targeting multiple phases of the viral replication cycle. Available anti-HIV drugs are distributed among five classes: nucleoside/nucleotide reverse transcriptase inhibitors (NRTIs); non-nucleoside reverse transcriptase inhibitors (nNRTIs); integrase inhibitors (INSTIs); protease inhibitors (PIs) and entry inhibitors, which include the fusion inhibitors (Fig. 1.14). In current cART regiment, at least three compounds are prescribed simultaneously, usually two NRTIs combined with a PI or an INSTI, to maximize the therapeutic effect of HIV inhibitors and avoid resistance development in response to monotherapy.

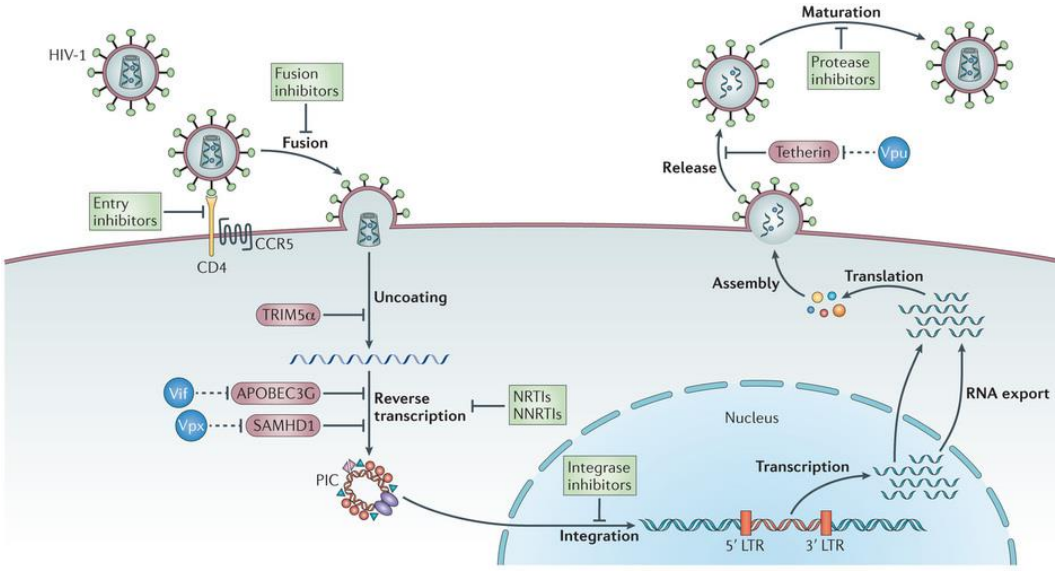


Figure 1.14: HIV inhibitors at viral replication cycle.
Adapted from Barre-Sinoussi.[143]

Antiretroviral therapy is effective in suppressing viral loads and transmission rates, enabling the reconstitution of $CD4^+$ cells count. However, this treatment fails to eradicate virus and consequently cure HIV-infected patients, implicating a lifetime treatment oftentimes associated with toxicity[144] and persistent immune activation that result in age-related and metabolic disorders.[145] On the other hand, the residual viremia present in cART-treated patients—highly derived from HIV latent reservoirs[146–149]— results in one of the most common causes for therapeutic failure of cART, the emergency of resistant strains. In most cases, latency occurs when infected $CD4^+$ T-cells return to a resting memory state, abolishing integrated provirus expression and establishing long-lasting HIV reservoirs.[146,147] Within these cellular reservoirs, HIV is 'hidden' from host immune system and antiretroviral agents, being responsible for viremia rebound and consequently de novo infection when latent cells are stimulated. The high mutational rate of HIV— related with error-prone nature of RT and virus rapid turnover—enables the selection of viral variants that escape the pressure of immune system and cART cocktail, providing them with a survival advantage. Polymorphisms among distinct HIV subtypes also influence virus susceptibility to inhibition and mutational resistance profile.[150] In other cases, cost and number/route of administrations could lead to discontinuous therapy and consequently sub-optimal concentration of cART and accelerated resistance arose. Other limitation of cART is related with differences in inhibitory capacity of distinct HIV types.[151] Only a few inhibitor classes such as NRTIs and INSTIs were described as active neutralizers of HIV-2 infection. Several PIs were potently compromised in inhibiting HIV-2, while nNRTIs and fusion inhibitors were in general ineffective. Specific

resistance to HIV-2 relates to differences between HIV types, limitations on HIV-2 clinical trials and design/selection of antiviral inhibitors almost exclusively based on HIV-1 clades and isolates. Finally, another factor contributing to limit cART efficiency could be the non-standard transmission route of HIV. Considering the consensual importance of cell-cell spread for HIV dissemination *in vivo*, activity of antiretroviral compounds against this transmission mode vs classical one has been compared. In general, cART efficacy is similar between cell-cell and cell-free virus spread, being the NRTI class of inhibitors the most affected by transmission switch.[152–154] Nonetheless, the massive concentration of viral particles in cellular contacts during cell-cell transmission could require a more efficient combination of highly active drugs. Cell-cell spread can also result in multiple integrated provirus and consequently originate novel recombination events and promote resistance strains appearance.[155,156] Regarding the entry inhibitors, and despite reported efficacy *in vitro*, more studies have to be conducted since endocytosis requirement/role in cell-cell spread remains unclear.[142,157,158]

In conclusion, and despite the increase in life expectancy and quality by cART, development of novel drugs continuous to be an urgent global need in HIV treatment. To reassure viral permanent and complete suppression, these therapeutic strategies have to be nontoxic, potent and effective against a broad spectrum of HIV subtypes as well as low-cost. One hypothesis consists in exploring novel classes of HIV neutralizers such as antibody, gene and RNA-based therapies against conserved/critical targets within virus replication cycle. A few ones are already under clinical trials development such as antibodies ibalizumab, an anti-CD4 entry inhibitor, and PRO-140 against CCR5 HIV co-receptor, zinc-finger/nuclease tool (SB-728-T) for knockout of *ccr5* gene or multiplexed RNAi molecules and alternative RNA-based agents (*ClinicalTrials.gov*; NHI USA).

1.3.1 Inhibition of HIV Entry

1.3.1.1 Entry inhibitors

The latest HIV inhibitors approved by USA Food and Drug Administration target the entry step of virus replication cycle, constituting the only class that comprises pre-infection drugs. This group has two members: T-20 against e-gp41 and maraviroc (MVC) targeting CCR5 chemokine receptor.

T-20 peptide has 36 aa based on HR2 sequence ($HIV_{643–678(LAI)}$) and blocks HR1-HR2 interaction during gp41-mediated fusion by competing with HR1 for binding.[159]

Therefore, T-20 is the only fusion inhibitor of HIV in the biopharmaceutical market. Despite well-documented T-20 potency against distinct HIV-1 clades and cell-cell spread,[160,161] issues have been associated with this inhibitor application in therapy. Resistance profile had emerged in treated patients, mainly related with mutations in a three amino acid sequence on HR1 region (GIV, 36-38 aa).[162,163] Additionally, T-20 is currently the unique anti-HIV drug that lacks oral administration, requiring to be subcutaneously injected twice daily. HIV-2 susceptibility to T-20 seems also remarkably reduced compared with HIV-1.[164,165] This last fact relates to several differences in structure, amino acid sequence and functionality of Env between HIV types. For example, HIV-2 entry process can be CD4-independent and mediated by alternative co-receptors, besides CCR5 and CXCR4, at least *in vitro*.[166,167] These complications, together with high production cost,[168] resulted in T-20 negligence on HIV therapeutic schemes.[169] In addition, this peptide has only been residually tested as microbicide and PrEP (pre-exposition prophylaxis) candidates, showing moderate results in preventing HIV transmission.

Second and third-generation alternatives to T-20 such as T-649, Tifuvirtide, T2635, Sifuvirtide and C34 have been developed, being these peptides also derived from amino acid sequences overlapping HR2 and adjacent motifs (Fig. 1.15).[170] However, and despite overall efficiency towards T-20-resistant strains, several mutations in HR1 have been described as decreasing or even abolishing HIV-1 susceptibility to these advanced HR2-based fusion inhibitors.[171]

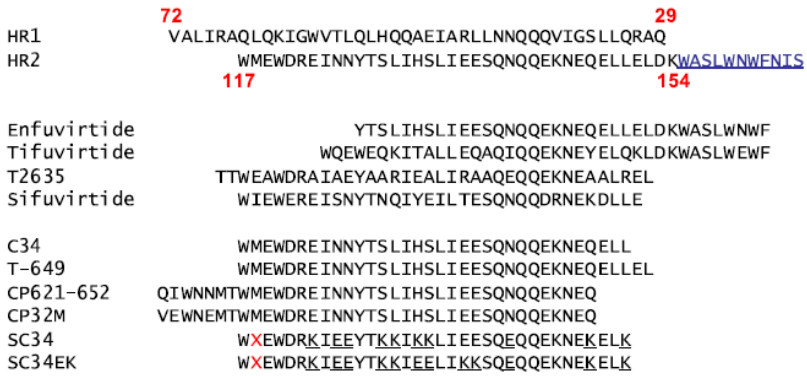


Figure 1.15: Pairing of HR1 and HR2 on post-fusion state of gp41 (Top) and amino acid sequences of T-20 (Enfuvirtide) and novel HR2-based fusion inhibitors (Bottom).
X represents norleucine. Adapted from Menéndez-Arias.[171]

MVC binds a hydrophobic pocket in transmembrane region of CCR5 similarly to other antagonists of this chemokine receptor.[172] MVC-CCR5 interaction results in conformational changes on receptor extracellular loops, blocking gp120 binding and consequently HIV fusion.[172] Similarly to T-20, some drawbacks[171] have been identified in MVC usage although reported efficacy in HIV treatment.[173,174] In consequence of restricted inhibitory activity towards CCR5-specific viruses, MVC can promote the outgrowth of CXCR4-tropic variants known to accelerate $CD4^+$ T cells depletion and consequently disease progression.[175] Several mutations on gp41 and gp120, not restricted to V3 loop, have also been associated with clinical resistance to MVC,[169] in addition to suggestion that higher doses of MVC are required to an efficient inhibition of HIV-2 isolates compared with variants of type 1.[164] Even though a few studies documented positive responses to MVC therapy in HIV-2 infected individuals, they were only conducted into 'salvage' therapeutic context.[176,177] Considering all limitations and similarly to T-20, MVC is not a first-line antiretroviral agent, being only applied in 'aggressive' therapeutic schemes for multidrug-resistant patients. Nonetheless, and despite recent failure in impairing SHIV infection of macaques,[178] MVC is the only entry inhibition under clinical development (Phase II) for HIV prevention after fully protected humanized mice from HIV-1 challenge.[179] MVC was also shown to protect humanized mice and macaques from SHIV infection when used as a microbicide, being currently tested under distinct formulations.[180–182]

Since T-20 and MVC approval, several entry inhibitors reached clinical trials, mainly CCR5 antagonists but also CXCR4 and CD4 blockers as well as g120 binders. Two mAbs are also being evaluated as entry inhibitor against CD4 receptor and CCR5 coreceptor. In resume, and despite advances so far, there is an urgent requirement for novel entry inhibitors.

1.3.1.2 Broadly Neutralizing Antibodies

Elicitation of broadly neutralizing antibodies (bNAbs) constitutes a huge hurdle in HIV treatment/prevention. Despite capability of majority HIV-infected individuals to produce strain-specific neutralizing antibodies a few months after infection, circulating viruses escape the immune response through a mutation-driven process and ensure ongoing infection. The ineffectiveness of immune response against HIV infection is mainly related to structural and chemical characteristics of Env (Fig. 1.16).[183] The outer domain of this viral glycoprotein presents antigenic diversity among HIV-1 clades[184,185]

and poor immunogenicity from extensive and shifting glycosylation. In contrast, conserved and critical-to-infection sites are conformational masked or steric occluded, being only transiently exposed during HIV fusion. Low density of functional Env complexes at virion surface combined with non-functional variants also contributes to HIV-immune escape. In a few HIV-infected individuals, the evolution of immune response though cycles of virus escape is believed to result in bNAbs production. By this way, these "elite controllers" (~1% seropositive individuals) are able to generate antibodies with *in vitro* cross-clade neutralization and aptitude to control virus replication and consequently delay AIDS onset in absence of cART.[186] Several bNAbs against distinct targets within Env glycoprotein are represented in Fig. 1.16.

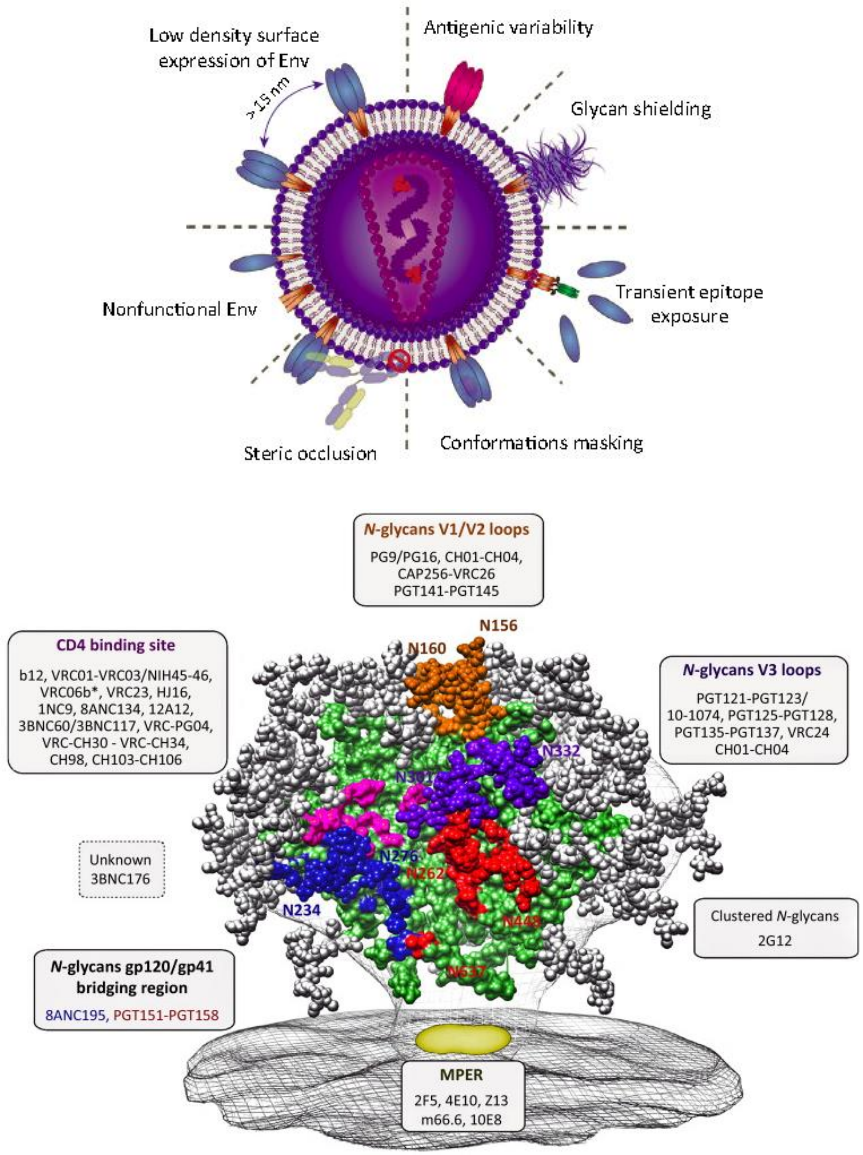


Figure 1.16: Immune-evasion characteristics (Top) and broadly neutralizing antibodies (bNAbs) of HIV (Bottom).

Top) Antigenic variability of Env-surface residues within an individual during HIV infection timeline and among HIV strains and clades; the extensive presence of glycans at Env-exposed regions, creating a "shield" against the antibody-mediated immune response; transient exposure of critical epitopes only during HIV fusion, for example membrane proximal region (MPER) on intermediate structure of gp41 and co-receptor binding site after gp120-CD4 engagement; conformational masking of, for example, CD4-binding site, sterical occlusion of critical sites to high molecular weight molecules such as antibodies; existence of non-function Env-complexes (non-cleaved Env precursors, gp41 six-helix bundles and gp120-gp41 monomers) and low density of Env molecules at virion surface. **Bottom)** Major targets of bNAbs on Env glycoprotein include CD4 binding site, V1/V2 loops, glycans at V3 loop, bridging region of gp120/gp41 interface region and membrane proximal external region (MPER) in gp41. Adapted from Mouquet.[183]

Overall, bNAbs can potently and broadly inhibit HIV infection *in vitro*, some of them with a neutralization potency (50% inhibitory concentration; IC50) of less than 1 µg/mL and a neutralization breadth of >90% such as targeting-gp41 10E8 and CD4bs-specific VRC01 and VRC02.[187] As an exception, some bNAbs such as 4E10 and 2F5 although presenting a potent and broad activity were reported as polyreactive, leading to an immune response against self and non-HIV antigens.[188,189] Surprisingly, polyreactivity was more observed in gp41-targeting bNAbs than gp120-specific ones.[190–194] On the other hand, rise of clinical resistance occurs more often for gp120-specific bNAbs. Regarding cell-cell spread, there is no consensus about bNAbs efficacy relative to this HIV transmission mode. While *in vitro* studies reported that only the most potent neutralizing antibodies, specific for CD4bs and V3 loop, efficiently inhibited cell-cell spread and even so with a concentration ten-fold higher than cell-free virus,[195] *in vivo* studies seem to indicate a similar potency against the two possible spread modes.[196–198] The exception are the gp41-MPER targeting bNAbs 4E10 and 2F5, which sensitivity seems decreased in simians infected with HIV/SIV chimeras (SHIV).[199] Of particular interest was the broadly CD4bs-specific VHHs isolated from llama,[200] namely the J3,[201] and the potent N6 bNAb[202]. The former is capable of potently inhibiting cell-cell spread *in vitro*,[203] including neutralizing antibodies-resistant pathway between macrophage and T-cells,[204] while the second one was reported to circumvent common resistance mechanisms of HIV virus. Additionally to reported vireamia suppression *in vivo*,[197,198,205] including human clinical trials,[206,207] recent studies revealed bNAbs protection of humanized mice and macaques challenged with HIV or SHIV chimeras, respectively.[208–217] Recently, Gautam et al. went further by demonstrating a long-lasting protection of several months, when macaques were repeatedly challenge with low doses of SHIV after single injection of a potent bNAbs cocktail.[218]

Although remarkably potency in therapeutically and prophylactically impairment of HIV infection, neutralization escape has already been described for some bNAbs.[205,216] Nevertheless, peculiar features of these looked-for antibodies such as high number of

somatic hypermutations, including rare nucleotide insertions and deletions, prevalence of kappa light-chains and extended CDR3 loops on heavy-chain domains (20-35 residues vs 14 residues)[183] can give insights into structure and sequence signatures for rational design of HIV vaccines and next-generation inhibitors.

An alternative scaffold to bNAbs is the dual-affinity re-targeting (DART) proteins, where the variable domains of distinct antigen-binding sites are incorporated into a disulphide-stabilized heterodimer. In the context of HIV therapeutics, diverse DART molecules, wherein one pair of variable domains recruits the cytotoxic T effector cells via CD3 receptor and the other one targets the HIV-1 Env on the surface of CD4-infected cells, have originated positive results in viremia suppression and partial restoration of immune system functionality.[219] Nonetheless, the safety and efficacy of this immune-inducible format have to be addressed in appropriate clinical trials. The recent advances on the engineering of chimeric antigen receptors (CARs) based on anti-HIV bNAbs also allowed to potentially target and kill infected cells despite absence of viral replication.[220,221] Furthermore, in one of the studies T $CD8^+$ cells harboring antiviral CARs were resistant to HIV infection due to *ccr5* disruption through gene editing.[220] These results contrast with the failed attempts on clinical applications of CARs for HIV treatment in the late 1990s using the extracellular domain of CDR4 receptor.[222] However, some clinical issues have been associated with CARs application such as the risk of toxicity[223] and immunogenicity induction as well as lack of CAR persistence to control the appearance of viral resistance in HIV context.

1.3.2 Inhibition of HIV Expression

1.3.2.1 Small Interfering RNAs

RNA interference (RNAi) is a post-transcriptional mechanism of gene silencing by which small RNA molecules, mainly small interfering and micro RNAs (siRNAs and miRNAs), induce degradation of target RNA messenger (mRNA).[224,225] While miRNAs have to be transcribed from cell genome as primary miRNAs (pri-miRNAs) and afterwards converted into precursor miRNAs (pre-miRNAs) to be nucleus exported, exogenous derived double-stranded RNAs (dsRNAs) can be directly loaded into RNAi pathway to be processed by Dicer into short duplexes of ~21 nucleotides with two-nucleotides 3'overhangs (siRNAs). These RNAi molecules are loaded into the RNA-induced silencing complex (RISC), where passenger strand is uncoiled and discarded,

enabling the guide strand to direct the complex towards homologous mRNA. RISC-endonuclease Argonaute arrests translation or cleaves target mRNA, leading to its consequent degradation. In contrast with miRNAs, siRNAs generally present full complementary with target mRNA and consequently single specificity (Fig. 1.17).[226,227] Mammalian cells employs RNAi pathway to regulate endogenous gene expression[228] and to degrade invading viral RNAs although this last role is still under debate depending on cell and virus types. Besides naturally occurring RNAi regulators, mimics of pri- and pre-miRNAs such as artificial miRNAs and short-hairpin RNAs (shRNAs) as well as synthetic dsRNAs and siRNAs can be engineered to trigger RNAi silence-specific pathway.

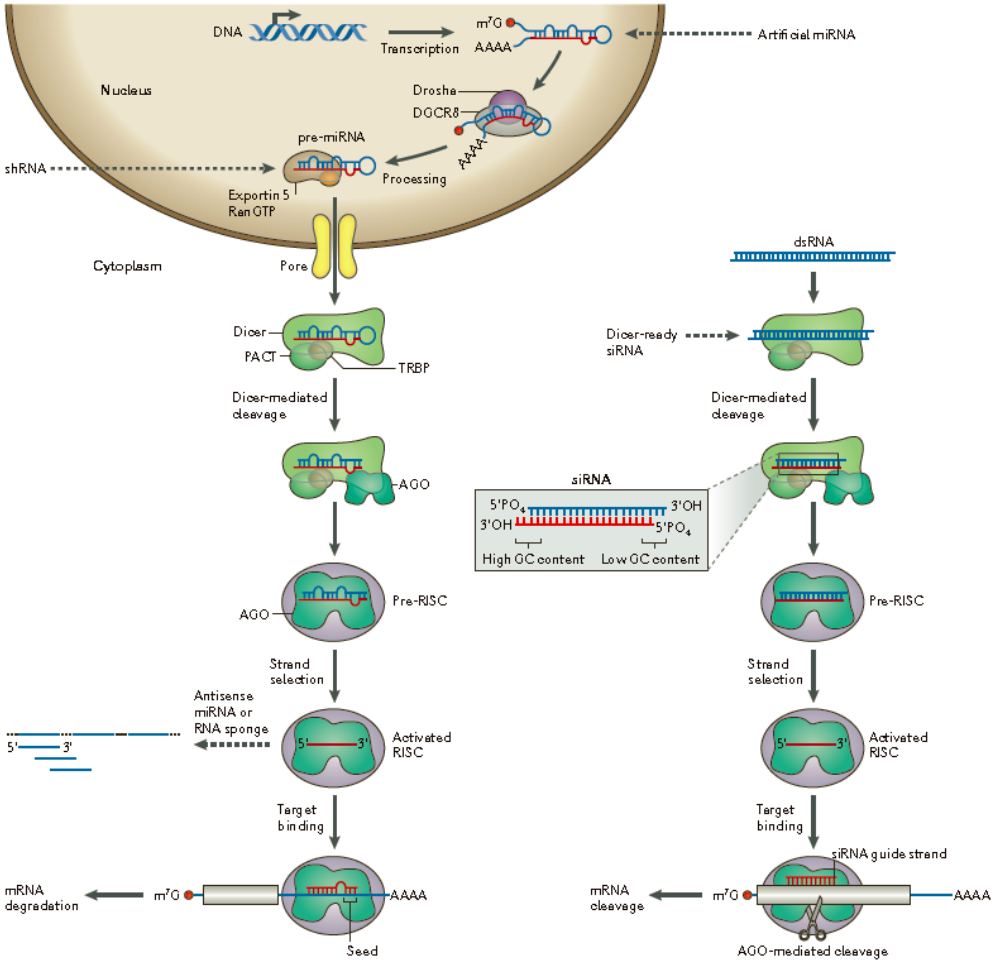


Figure 1.17: RNA interference mechanism (RNAi) in mammalians.
Left) Primary miRNA (pri-miRNA) is transcribed from cell genome and recognized by protein complex: Drosha and DGCR8. Drosha processes the pri-miRNA into the stem-loop precursor miRNA (pre-miRNA), which is transported by Exportin 5 from nucleus to cell cytoplasm. Pre-miRNA is cleaved by Dicer into mature miRNAs. Depending on miRNA complementarity with target mRNA, this can be degraded or its translation repressed. **Right)** Long double-stranded RNAs (dsRNAs) are processed into siRNAs by Dicer-R2D2 heterodimer, components of the RISC-loading complex. Processed RNA molecule is loaded into RNA-induced silencing complex (RISC), where passenger RNA (sense strand) is uncoiled and discarded by endoribonucleases Argonaute 2 (AGO2) and component 3 promoter of RISC (C3PO). The

remaining anti-sense strand guides RISC to the complementary target mRNA, which is cleaved by the AGO2 endonuclease activity and subsequently degraded by exonucleases. Adapted from Davidson & McCray.[229]

Triggering gene silencing exogenously through the usage of tailor-made molecules constituted a breakthrough in therapeutics.[225] Despite origin, chemical and processing differences, all RNAi regulators present broad activity translated into virtually silencing of any interest gene, namely 'non-druggable' targets such as proteins lacking function or accessible conformation to conventional therapeutics.[230] Nevertheless, synthetic siRNAs are expression-independent from DNA vectors and surpass the requirement for nucleus exportation and Dicer processing step, being directly loaded into RISC, even though Dicer-independent shRNAs had already been designed.[231] Accordingly, siRNAs are the most represented class of RNAi molecules in human clinical trials with activity in genetic and infectious disorders, cancer and ophthalmic conditions.[232] Nonetheless, limitations continue to hold back RNAi application in therapeutics.

Besides siRNA sensitivity to serum nucleases, off-target silencing[233] although not yet documented *in vivo*, and activation of immune system,[234] the lack of a specific and efficient delivery system continues to be the major challenge in RNAi therapeutics. With exception of particular cell types such as pulmonary epithelial cells that uptake siRNAs from surrounded environment,[235] intrinsic characteristics of these minimal RNA molecules such as hydrophilicity and anionic nature are responsible for low cellular uptake. Moreover, siRNAs are too large to diffuse across membranes, but too small to be maintained in circulation, being rapidly cleared by kidney excretion.[230] Even though chemical modifications have increased siRNAs half-life,[236,237] this hurdle continues to be a major contributor to membrane-impermeable feature of these small RNA-based molecules.

1.3.2.1.1 Strategies for Delivery of Small Interfering RNAs

For some indications, local administration of siRNAs is sufficient to ensure a therapeutic effect as reported previously for intranasal, intraocular, intratumoral, intramuscular and intracerebral approaches *in vivo*.[230] Nonetheless, systemic delivery is required to hamper some disease types, where target cells/tissues are inaccessible or widespread though human body.

The vehicles for systemic delivery of siRNAs can be divided into physical methods, viral vectors and non-virus based approaches. These systems can furthermore be classified into specific and non-specific methods. Physical vehicles of siRNA delivery include high-pressure injection[238], electroporation[239,240], ultrasound,[241] and gene gun[242]. Although positive results in mice, safety issues limit the therapeutic usage of these methods. For example, high-pressure injection was reported to cause right-sided heart failure. An efficient alternative for the delivery of RNAi molecules are the virus-like particles such as adeno-associated virus (AAV), adenovirus, herpes virus and retrovirus (including lentivirus). However, ethical and safety concerns continue to retreat viral delivery methods, mainly the need for chromosomal integration of retroviral vectors—the most common approach for stable expression of shRNAs. Viral vectors are also time-consuming to produce, size-constrained in the case of AAVs and limited to deliver shRNAs and artificial miRNAs, which overexpression could cause cellular toxicity[243,244] and usage cannot be discontinued or altered in response to treatment evolution, disorder type or population heterogeneity. Non-viral methods represent the most advanced category in systemic delivery of siRNA molecules and the unique that include cell-type targeting.

Specificity, namely through cell-surface receptor delivery, results in an efficiency increase and, simultaneously, side effects reduction of therapeutic siRNAs. These molecules can be chemical conjugated to cell-receptor ligands of non-antibody type or fused to antibody-based molecules targeting the cell surface.[245] Common ligands of non-antibody type include cholesterol,[236] folic acid[246] and aptamers.[247–249] Regarding antibody-mediated delivery, Fabs and scFvs are among the most used targeting-entities[250–252] to avoid immune stimulation prompted by Fc region. siRNAs molecules can also be complexed with cell-penetrating peptides[253] such as arginine-rich protamine and poly-arginine-peptides[254] or polymeric[255] and lipidic-based[256,257] nanocarriers, which in turn can be decorated with targeting molecules such as antibodies. Surprisingly, RNAi molecules were also described to move between mammalian cells through connexin-specific gap junctions, after donor cells were transfected with shRNA.[258] Bacterial-derived cells with no genome (minicells) were as well used to express or carry RNAi triggers.[259] Delivery of RNAi molecules could be performed *ex vivo*, where target cells are isolated, treated with RNA-based molecule, and re-injected into the patient. Some non-based virus methods are represented in Fig. 1.18.

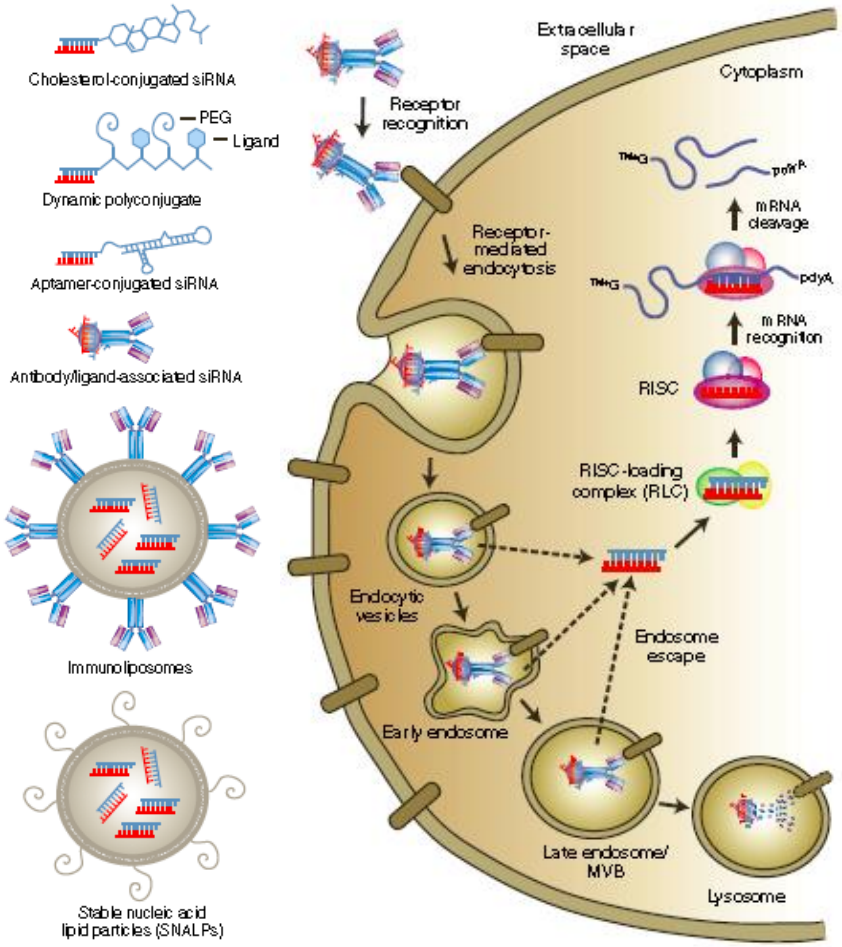


Figure 1.18: Non-based virus delivery systems and cell-surface receptor delivery pathway of siRNAs. **(Left).** Cell-specific delivery of siRNAs can be mediated by cholesterol and other ligands of cell-surface receptors. Cell surface-targeting aptamers and antibodies can also be fused with siRNAs. These RNAi molecules can also be encapsulated in lipid nanoparticles decorated or not with cell-surface antibodies (SNALPs and immunolipossomes, respectively). **(Right).** Antibody recognizes and binds a cell-surface receptor, which internalizes through endocytosis, siRNA escape from endossome and is loaded into RISC-loading complex. Adapted from Dominska & Dykxhoorn.[260]

Even though there is no 'ideal' method for *in vivo* siRNA delivery, since it depends on targeting cell-type, location and desirable therapeutic effect, antibody-mediated delivery seems to gain ground relative to others delivery approaches, mainly to liposome-based methods known to easily trigger diverse inflammatory pathways[261] and difficult intracellular release of siRNA.[256]

1.3.2.1.2 Small Interfering RNAs in HIV Therapeutics

A diverse array of therapeutic siRNAs targeting key phases of HIV infection has been developed (database http://crdd.osdd.net/raghava/hivsir/),[262] being conservation of RNAi target crucial to avoid clinical resistance. RNAi therapeutics theoretically allows to target any sequence within HIV genome, which clearly exceeds cART capability. Moreover,

potency of RNAi, mainly when distinct regulators are multiplexed as a single therapy, can decrease dosage requirement and consequently possible off-target effects.

Several RNAi triggers were reported to efficiently suppress HIV in cell lines as well as primary T-cells and macrophages. However, the lack of an efficient and cell-specific delivery system impairs a potent therapeutic effect *in vivo*. This fact is even more dramatic in the context of HIV infection because of T-cells known resistance to lipid-based transfection. Accordingly, viral vectors constitute the prime choice for delivery of RNAi-based therapeutics in HIV and the unique strategy recently on clinical trials development, but only as *ex vivo* approaches.[261,263] In DiGiusto et al.,[264] haematopoietic progenitor cells isolated from peripheral blood were transduced with a single lentiviral vector expressing an anti-*tat/rev* shRNA, a CCR5-targeting ribozyme and a nucleolar localizing TAR decoy and afterwards re-injected into patients with AIDS-related lymphoma. In addition to *tat/rev* mRNA degradation induced by shRNA, catalytic ribozyme cleaves CCR5 mRNA, while TAR decoy sequesters Tat from its cognate target on HIV promoter. On the other hand, Symonds and co-workers[265] developed a self-inactivating lentiviral vector codifying an anti-CCR5 shRNA and the fusion inhibitor peptide C46, which were transduced in hematopoietic stem/progenitor and $CD4^+$ cells. Like cART, a combinatorial approach with multiple and distinct targeting molecules is expected to prevent emergency of RNAi-resistance virus and to compromise viral fitness of possible escape variants. Nevertheless, safety issues resulting from expression, integration and induction of interferon responses along with difficulties in maintaining stable transgene expression in progeny T-cells for impairment of HIV resistance have been attributed to shRNA-based delivery methods.[266–272] Moreover, an shRNA-based therapeutic scheme is unable to be adapted in response to escape viral mutations. Consequently, diverse strategies were already tested for the delivery of anti-HIV siRNAs, where antibody-mediated approaches stand out among these.

Regarding antibody-mediated delivery, Kumar et al. successfully delivered anti-HIV siRNAs to humanized mice through targeting T-cells CD7 marker with a scFv-9arginine peptide.[273] Previously, Song et al. had already fused a Fab antibody fragment with RNA-binding protamine to specifically deliver a siRNA against *tat* mRNA through viral gp120 glycoprotein. Additionally, this fusion protein suppressed tumour growth when loaded with anti-oncogenic siRNA and injected into mice harbouring Env-expressing tumour.[250] Likewise, others targeted gp120 to deliver anti-HIV siRNAs, this time with aptamer-

siRNA chimeras that suppressed viral load and restored $CD4^+$ count.[274,275] CD4 aptamer-siRNA chimeras were also reported to protect humanized mice from HIV sexual transmission, when applied as microbicide.[276,277] Immunoliposomes were as well described as effective in the delivery of therapeutic siRNAs to HIV prime targets in humanized mice.[278] Some non-targeting strategies have been under development for anti-HIV siRNA delivery such as electroporation and polymer, lipid and dendrimer nanoparticles.[261] All these methods present advantages and disadvantages. For example, antibodies and aptamers can residually protect siRNAs from serum nucleases, but provide an efficient and targeted delivery, diminishing side-effects occurrence. On the other hand, aptamer-siRNA chimeras lack of immune activation is counteracted by its rapidly clearance though kidney excretion.[279] On the contrary, antibody-siRNA constructions may trigger host immune system, but they are maintained long time in circulation. Targeted nanocarriers such as immunoliposomes can increase siRNA stability in serum but at the same time may stimulate a strong immune response. Non-targeting strategies, in addition to lack of specificity and consequently a decrease in efficiency, have tendency to induce interferon responses, mainly lipid nanoparticles, and in case of naked siRNAs present short half-life and bloodstream instability.

1.3.2.2 Artificial Transcription Factors Based on Zinc-finger Proteins

Zinc-finger proteins (ZFPs) comprise one of the most important and frequently codified class on eukaryotic genomes. ZFPs are constituted by tandem repeats of small motifs, the zinc-finger domains, being each one stabilized by hydrophobic interactions and the coordination of a Zn^{2+} ion. These peculiar proteins also present a tremendous diversity in structure and function, capable of interacting with diverse biological macromolecules including DNA, RNA, proteins, and lipids. Specific-nature of DNA-binding capacity is one of the most interesting and characterized feature of ZFPs, mainly Cys_2His_2 family. Cys_2His_2 domains have normally ~30 aa in length and present a characteristic ββα structure comprising two conserved cysteines and histidines residues, which are responsible for zinc binding and consequently ZF stabilization (Fig. 1.19).[280]

Zinc-finger domains recognize 3-4 bp of target DNA duplex. Key positions -1, 2, 3 and 6 on helix portion, known as recognition helix, are the main determinants of DNA-binding specificity (Fig. 1.19)[280–283] although the importance of a fifth nucleotide-residue contact was recently hypothesized.[284] Nonetheless, non-canonical interactions of these

positions and the role of flanking residues should not be unvalued, forward requiring the confirmation of designed ZFPs specificity for target DNA.[285,286] Individual zinc-finger domains with distinct specificities can be covalently linked in tandem repeats (polydactly ZFPs) to recognize extended stretches of DNA, theoretically targeting any sequence of interest.[287] Liu et al. was the first to validate the targeting of 18 bp DNA stretches, generally unique in human genome, by assembling two three-fingered proteins via ZFPs-naturally occurring linker TGEKP[288] although non-canonical linkers have also been used.[289–291] Recently, DNA-binding domains covering all possible 64 combinations of genetic code were identified through systematic screenings of large synthetic ZF libraries.[284] Together with the already available zinc-finger sets from Barbas laboratory,[292] Sangamo BioSciences[293] and ToolGen[294] companies, this discovery extends ZFPs application to a brand new level.

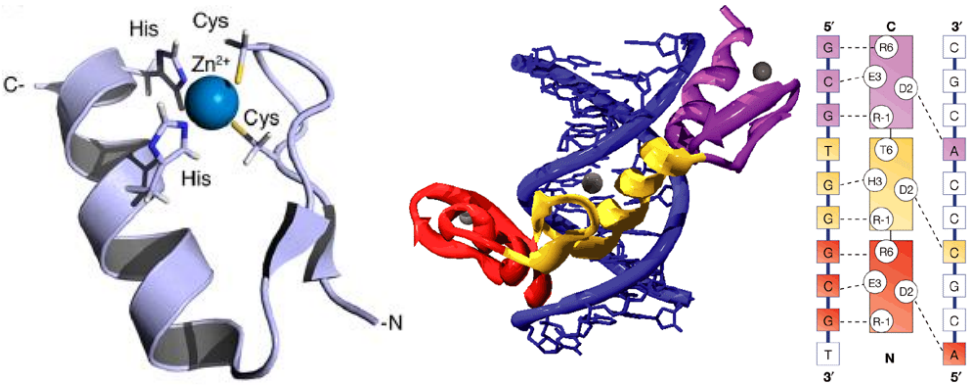


Figure 1.19: Structure of a zinc-finger domain and interactions between zinc-finger protein and DNA. Structure of ZFP-DNA complex **(Left)**. The zinc-finger protein with three zinc-fingers 1, 2 and 3 (red, yellow and violet) binds the major groove of DNA (blue). The grey spheres represent the Zn^{2+} ions. Sequence-specific binding of DNA-ZFP **(Right)**. The recognition helices are represented as squares in diagram centre. Key residues for ZF-DNA binding (positions -1; 2; 3 and 6 in surface of ZF α-helix) are represented by single-letter code. Dashed lines represent the interactions between zinc-fingers and the two strands of DNA. ZFP binds three contiguous nucleotides of DNA in 3'-5' orientation, with a possible forth interaction with the adjacent strand. Adapted from Gaj et al.[295] and Jamieson et al.[296]

Individual zinc-finger domains can be selected from combinatorial libraries, elected among the natural available repertoire or rationally designed, being afterwards modularly assembled towards desirable DNA sequence.[296–298] Alternative engineering methods for ZFPs design manage the context-dependent interactions by considering the entire sequence of target DNA instead of each triplet individually. Some examples are the bipartite,[299] sequential,[300] bacterial "two-hybrid"[301] and oligomerized pool engineering[302] methods. In general, these second-generation strategies improve quality of resulting DNA-binding proteins although cost and complexity are oftentimes increased. To make the most of customized ZFPs, these multimers of DNA-binding domains can be *in vitro* associated with functional moieties such as nucleases, recombinases, transposases,

methyltransferases, integrases and transcription factors (Figure 1.20).[287,303–309] One major application of synthetic zinc-fingers is on modulation of gene expression. Artificial transcription factors (ATFs) activate or repress target genes depending on engineered fused domain. For example, six-fingered proteins fused with the tetrameric repeat of herpes simplex VP16 activation domain (VP64) up-regulated expression of *epidermal growth factor receptor 2 and 3* genes, while fusion with Krüppel-associated box (KRAB) resulted in their repression.[310] ZFPs alone can also repress gene expression by sterically occluding transcription elongation or initiation and co-activators binding.[308] Additionally, site-specific methylation mediated by ZFP-methyltransferase can trigger gene silencing pathways.[311] The efficacy of ZF-based artificial regulators in activating or repressing gene expression have been validated in several *in vitro*[288,306,312,313] and *in vivo*[314,315] models.

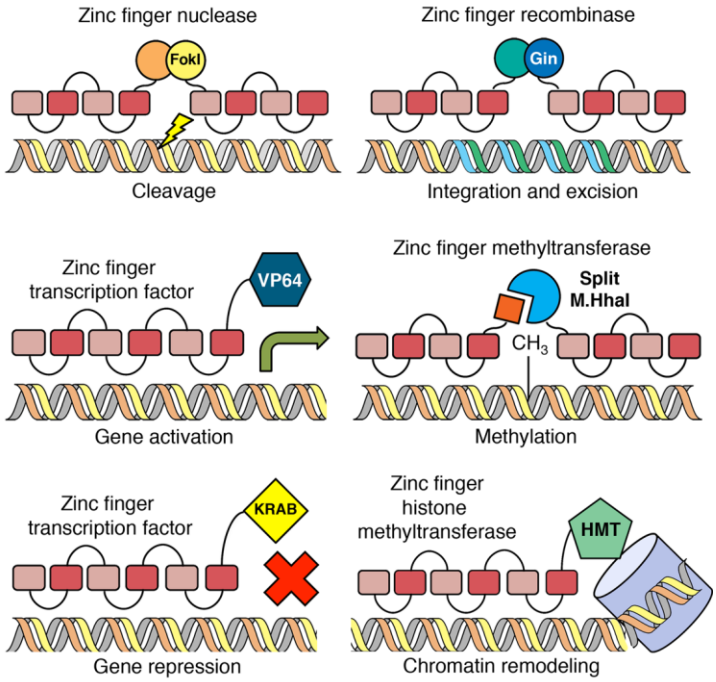


Figure 1.20: Applications of zinc-finger proteins in genome engineering.
Zinc-finger proteins can be fused or linked to nucleases such as FoxI to cleavage target DNA, recombinases, transcription factors such as VP64 activator or KRAB repressor to regulate gene expression, and methyltransferases. Adapted from Gersbach et al.[292]

1.3.2.2.1 Strategies for Delivery of Zinc-finger Transcription Factors

Virus-like particles (limitations described in section 1.3.2.1.1 Strategies for delivery of small interfering RNAs) together with toxicity-related[316,317] transfection and electroporation methods[318] have been the most common carriers of ZFPs[319–322] until a breakthrough report from Gaj *et al.* described zinc-fingers as naturally cell-penetrating

domains.[323] Moreover, the same authors reported efficient ZFPs-mediated delivery of functional large proteins into cytoplasm of diverse mammalian cell types.[295] Recently, a ZF-based ATF was even capable of cross the blood-brain barrier and restore the expression of a brain endogenous gene, when fused with the well-known HIV Tat cell-penetrating peptide.[324] Accordingly, Liu et al. had previously observed that cell-permeability of zinc-fingers was enhanced by incorporation of successive nucleolar localization signals (NLS)—highly positively charged peptide known to innately cross cell-membrane.[325] Nevertheless, some laboratories are still reporting virus-based approaches to successfully deliver zinc-finger based effector molecules, mainly nucleases.[326–328] These findings, together with the fact that zinc-fingers constitutes currently the most represented genome engineering tool under clinical trials development (NCT0047693),[329] enhanced the powerful application of these zinc coordinated domains.

1.3.2.2.2 Zinc-finger Transcription Factors in HIV Therapeutics

ATFs assume a significant role in HIV inhibition, since genome editing therapeutic efforts have proven difficult to overcome the rapidly viral escape provided by high mutation rate of RT or could even contribute for acceleration of resistance development.[330,331] Although the importance of transcription activators in hijacking HIV latency, this thesis will focus on role of repressor ATFs for the inhibition of HIV transcription and replication. Several repressor ATFs were already developed to shutdown LTR promoter and therefore prevent viral proteins transcription and completion of HIV replication cycle.[332,333] One of the most potent antiviral ATFs described was KRAB-HLTR3,[333] which presents the well-known repression domain KRAB[334,335] as effector molecule. KRAB seems to present the most potent transcriptional repressor activity, when compared with other repressor domains such as ERF repressor domain or mSin3 interaction domain.[310,333] KRAB-HLTR3 could inhibit HIV replication in cell lines as well as primary cells with potent activity. This potency can also be explained by the overlapping of HLTR3-binding site and two SP1-binding sites on HIV promoter, leading to the competition between KRAB-HLTR3 and SP1 transcription activator for LTR binding. Another ATF, designed to target the highly conserved tRNA primer-binding site (PBS) on LTR promoter, reported to inhibit HIV replication in primary T-cells.[332] Other ATFs were developed to target PBS region and SP1-binding sites on HIV promoter, resulting in the inhibition of viral transcription and replication.[336,337] In all these reports, transfection and lentiviral methods delivered the anti-HIV ATFs.

There are currently no ATFs under clinical trials development against HIV infection, only a zinc-finger nuclease targeting *CCR5* gene is currently being tested as an ex vivo approach. In this strategy, autologous CD4+ T-cells modified with viral vector codifying therapeutic zinc-finger-nuclease gain resistance to HIV R5 infection.[338,339]

In general, lack of potency and breath as well as specificity in targeting/delivery of HIV therapeutic molecules continue to be major hurdles in treatment/eradication of AIDS disease. In this thesis, we propose to contribute for the closing of these gaps through the engineering of innovative antibody-based molecules.

CHAPTER II

Development of synthetic light-chain antibodies as novel and potent HIV fusion inhibitors

Catarina Cunha-Santos[1], Tiago N. Figueira[2], Pedro Borrego[1,3], Soraia S. Oliveira[1], Cheila Rocha[1,3], Andreia Couto[1], Cátia Cantante[1], Quirina Santos-Costa[1], José M. Azevedo-Pereira[1], Carlos M.G.A. Fontes[4], Nuno Taveira[1,3], Frederico Aires-Da-Silva[4,5], Miguel A.R.B. Castanho[2], Ana Salomé Veiga[2], and Joao Goncalves[1]

[1] *Research Institute for Medicines (iMed.ULisboa), Faculty of Pharmacy, Universidade de Lisboa, Lisbon, Portugal*

[2]*Instituto de Medicina Molecular, School of Medicine, University of Lisbon, Lisbon, Portugal*

[3]*ISCSEM-Centro de Investigação Interdisciplinar Egas Moniz, Instituto Superior de Ciências da Saúde Egas Moniz, Monte de Caparica, Portugal*

[4]*CIISA-Interdisciplinary Centre of Research in Animal Health, Faculdade de Medicina Veterinária, Universidade de Lisboa, Lisboa, Portugal*

[5]*Scientific BU Manager of TechnoAntibodies*

AIDS. 2016 Jul;30(11):1691-701

2.1 Abstract

Objective: To develop a novel and potent fusion inhibitor of HIV infection based on a rational strategy for synthetic antibody library construction.

Design: The reduced molecular weight of single-domain antibodies (sdAbs) allows targeting of cryptic epitopes, the most conserved and critical ones in the context of HIV entry. Heavy-chain sdAbs from camelids are particularly suited for this type of epitope recognition because of the presence of long and flexible antigen-binding regions [complementary-determining regions (CDRs)].

Methods: We translated camelid CDR features to a rabbit light-chain variable domain (VL) and constructed a library of minimal antibody fragments with elongated CDRs. Additionally to elongation, CDRs' variability was restricted to binding favorable amino acids to potentiate the selection of high-affinity sdAbs. The synthetic library was screened against a conserved, hidden, and crucial-to-fusion sequence on the heptad-repeat 1 (HR1) region of the HIV-1 envelope glycoprotein.

Results: Two anti-HR1 VLs, named F63 and D104, strongly inhibited laboratory adapted HIV-1 infectivity. F63 also inhibited infectivity of HIV-1 and HIV-2 primary isolates similarly to the US Food and Drug Administration-approved fusion inhibitor T-20 and HIV-1 strains resistant to T-20. Moreover, epitope mapping of F63 revealed a novel target sequence within the highly conserved hydrophobic pocket of HR1. F63 was also capable of interacting with viral and cell lipid membrane models, a property previously associated with T-20's inhibitory mechanism.

Conclusion: In summary, to our best knowledge, we developed the first potent and broad VL sdAb fusion inhibitor of HIV infection. Our study also gives insights into engineering strategies that could be explored to enhance the development of antiviral drugs.

2.2 Introduction

Significant advances in antiretroviral therapy have occurred as the approval of the first fusion inhibitor, T-20.[340] T-20 peptide derives from the C-terminal region (HR2) of gp41 fusion protein from $HIV_{643-678(LAI)}$.[341,342] By competitively binding the N-terminal region (HR1), T-20 impairs HR1-HR2 interaction[341] and consequently the formation of the six-helix bundle (6HB) structure, responsible for HIV fusion (reviewed in Wilen et al.[122]). Despite the well-characterized antiviral potency of T-20, clinical resistance has been reported in HIV-1-infected patients.[163] Additionally, T-**20** is described as antigenic,[343] highly expensive, protease-susceptible (no oral administration) and "*pharmacokinetically limited"***,** among other limitations.[168] Overall, development of novel HIV fusion inhibitors with improved biophysical and pharmacokinetic properties is required.

Antibody fragments emerged to overcome issues associated with high-molecular weight of native antibody structure (IgG), mainly the targeting of cryptic epitopes and the penetration into densely packed tissues. Single-domain antibody (sdAb) is currently the smallest functional antibody fragment, only constituted by the antibody heavy-chain or light-chain variable domains (VH or VL).[15] Additionally to the reduced size, complementary-determining regions (CDRs; antigen-binding regions) of sdAbs can be easily engineered to develop specific and high-affinity binders. sdAbs also present excellent biophysical properties such as high stability, solubility and low toxicity.[344] Despite these beneficial features, only the therapeutic potential of VH domains have been intensively explored.[345] Nevertheless, several reports have demonstrated that VL domains present excellent biophysical properties such as high expression yield, resistance to aggregation and proteases, stability and high reversibility of thermal unfolding, in some cases better than VHs.[47,48,63,69] Moreover, the stability of VL domains was further evidenced by the proved functionality of these sdAbs in the absence of disulfide bonds[46,64] or in the reducing cellular environment[65,66].

Here, we engineered a VL sdAb with elongated CDRs that broadly and potently inhibits HIV-1 infection by targeting a well conserved and crucial-to-fusion sequence on HR1. Despite the clinical resistance to HR1-targeting T-20, this region contains highly conserved residues among HIV-1 subtypes and isolates,[116] representing a major target to HIV infection impairment. Anti-HR1 VLs were selected by phage display technology

from a restricted combinatorial library. Epitope mapping of the two most potent antiviral VLs—selected against an HIV-1 laboratory-adapted strain—showed that these inhibitors target a highly conserved and critic region within HR1. One VL (F63) showed high potency to inhibit HIV-1 and HIV-2 primary isolates with comparable T-20 activity. For last, we demonstrated that F63 also interacts with lipid membranes, a key ability of potent HIV entry inhibitors[346,347] that correlates with their mechanism of action.

2.3 Materials and Methods

2.3.1 N36 antigen

N36 was synthetized (Thermo Fisher Scientific) based on HIV-1 strain $HXB2_{Env546-581}$ and solubilized in 40% (v/v) DMSO at 1 mg mL^{-1}.

2.3.2 VL construction plasmids

Modified pT7-FLAG-2 plasmid (Sigma-Aldrich) with the cloning sequence from pComb3X plasmid (ompA peptide leader to amber stop codon) inserted into HindIII/BglII sites was obtained from Technophage, Lisboa, Portugal.

The anti-HEL VLs were generated by PCR through the CDR3 grafting of a stable rabbit dAb[348] derived from a scFv specific for HIV-1 Vif protein[349]. VL sequence used as template was the following: 5'-GAGCTCGTGCTGACCCAGACTCCATCCTCCGTGTCTGCAGCTGTGGGGGGCACAGTCACC
ATCAATTGCCAGGCCAGTCAAAGTGTTTATAATAACAACAACTTAGCCTGGTATCAGCAGAAACCAGGGCAGCGTCCCAAGCTCCTGATCTATGGTGCATCCGATCTGGCATCTGGGGTCTCATCGCGGTTCAAAGGCAGTGGATCTGGGACACAGTTCACTCTCACCATCAGCGGCGTGCAGTGTGCCGATGCTGCCACTTACTACTGTCAAGGCGAATTCAGTTGTGTTGGTGGTGATTGTTTTGCTTTCGGCGGAGGGACCGAGCTGGAGATCCTA-3'. Each VL construction was originated by assembling of two purified PCR fragments (example: CDR3 22 aa F fragment x CDR3 22 aa R fragment; Table 2.1) by PCR overlap. PCR-overlap fragments were inserted into the SfiI sites of modified pT7-FLAG-2. VL sequences were verified by DNA sequencing analysis.

Anti-HR1 VLs were cloned into the NheI/XhoI sites of pET-28a (+) (Novagen, Merck Millipore; primers sequence in Table 2.1 Supplementary Information). F63 dimer with a medium linker (ML, $[GGGS]_2$) was constructed by a restriction/ligation strategy

previously described in Oliveira *et al.*[350] Briefly, two PCR fragments were originated (F63 monomer N-terminal and F63 monomer C-terminal, Table 2.1 Supplementary Information), digested by NotI and assembled together. The resulting fragment was inserted into the Nhe/XhoI sites of pET28a (+).

2.3.3 ELISA measurements

ELISA assays were performed as described.[43] Briefly, high-binding plates (Costar 3690, Thermo Fisher Scientific) were coated with 1 µg of hen egg-white lysozyme (HEL; Sigma Aldrich)/bovine serum albumin (BSA) or high-adsorption peptide plates (Immobilizer-amino plates, Nunc, Thermo Fisher Scientific) were coated with 500 ng of N36/BSA according to manufacturer's instructions. After blocking with PBS-BSA, 500 ng of purified VLs were added to the wells and further incubated. VLs were detected by a HRP-conjugated anti-HA monoclonal antibody (anti-HA-HRP; Roche Diagnostics). For the competitive ELISA, 30 pmol of VL constructions were incubated with different quantities of HEL or N36 peptide in solution at 37 ºC. After 1 h, this mixture was added to HEL or N36-coated wells and further incubated. For epitope mapping, immobilizer-amino plates were coated with 1 µg of each HR1 peptide from HIV-1 Clade B (MN) Env Peptide Set (NIH AIDS Reagent Program)/BSA according to manufacturer's instructions. After BSA-PBS blocking, 100 ng of purified VLs were added to the wells and further incubated during 1 h at 25 °C.

2.3.4 Library construction and selection

Synthetic library was generated using the VL derived from the anti-Vif scFv[349] as scaffold and accommodating a maximum of 22 aa in CDR1 and CDR3. Briefly, CDR1 randomization was first introduced by two PCR amplifications (primers sequences in Table 2.1 Supplementary Information), using the DVN degenerate codon. D codifies for bases A, G or T; V codifies for bases A, C or G and N codifies for bases A, C, G and T. The amino acid distribution for the DVN codon can be found in CodonCalculator algorithm (http://guinevere.otago.ac.nz/cgi-bin/aef/CodonCalculator.pl).[351] The two PCR fragments were subjected to PCR overlap (primer sequences in Table 2.1 Supplementary Information). The resulting PCR product was then submitted to a second round of amplification to introduce the randomization in CDR3 (primers sequences in Table 2.1 Supplementary Information). The final PCR product was cloned into the SfiI sites of pComb3X phagemid (provided by C. Barbas III), designed to display recombinant proteins on the surface of filamentous phage M13 in a monovalent format [352]. The

synthetic library represented by ~$8.0x10^9$ independent clones was transformed into the amber-stop codon suppressor strain *Escherichia coli* (*E. coli*) K12 ER2738 (New England Biolabs) that allows the expression of the synthetic VLs in fusion with phage p3 surface-protein.

Selection of VLs against HR1 peptide was performed by phage display technology, as previously described [353]. Briefly, phages from the library were cycled through three rounds of binding selection (pannings) against HR1-coated 96-well Immobilizer-Amino Plates (Nunc, Thermo Fisher Scientific). Adsorption of HR1 peptide was performed according to manufacturer's instructions, using decreasing quantities of antigen throughout the pannings (1000 ng in the first, 750 ng in the second and 500 ng in the third panning). Phages expressing the recombinant VLs were incubated with HR1 peptide during 2 h. Washes were performed five times with PBS-Tween 20 0.05% (v/v). Bound phages were eluted with trypsin for 30 min and then amplified in *E. coli* K12 ER2738 with the addition of VCSM13 helper phages (Agilent Technologies). The third panning-selected VLs pool was cloned into previous inserted SfiI sites of pT7-FLAG-2 expression vector and transformed into *E. coli* BL21 (DE3) (Invitrogen, Thermo Fisher Scientific). For protein expression, individual clones from the third-panning were grown in round-bottom 96-well plates using the Overnight Express Auto-induction System 1 (Merck Millipore) according to manufacturer's instructions. After overnight expression, cells were lysed with BugBuster Protein Extraction Reagent (Merck Millipore) and cell extracts harvested through centrifugation. ELISA screening was performed as described above, using 500 ng of HR1 peptide as antigen. The five clones that presented more binding to HR1 peptide over BSA were subjected to DNA sequence analysis.

To the anti-HIV screening, 1 µg of bacterial extract from the five anti-HR1 VLs selected by phage display and ELISA were incubated with 40 000 infectious units (IU) of HIV-1_{NL4-3}. After 3 h at 37 °C, 40 000 TZM-bl cells were infected with the HIV/inhibitors mixture. The medium was changed 3 h after infection. At 48 h post-infection, *β-galactosidase* expression was quantified as described.[43]

2.3.5 VLs expression and purification

$VL_{parental}$ was provided by Technophage, Lisbon, Portugal. Anti-HEL VL constructions were transformed into *E. coli* strain BL21 (DE3) cells and expressed to the bacterial periplasmic space due to the presence of OmpA leader from pComb3X plasmid. VLs were growth in 0.5 L of super broth (SB) medium supplemented with 100 µg mL^{-1}

ampicillin and 20 mM $MgCl_2$ at 37 °C until OD_{600} of 0.9. Protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and lasted for 16 h at 30 ºC. To evaluate VLs expression profile, 10 mL of each VL culture was recovered at different time points (0, 1, 2, 6 and 16 h). Bacterial pellets were harvested by centrifugation and resuspended in 20 mM NaH_2PO_4 pH 7.4, 500 mM NaCl and 20 mM imidazole for purification or PBS for western-blot supplemented with a protease inhibitors cocktail (Roche Diagnostics). Cells were lysed by sonication. Soluble protein was purified by immobilized metal ion affinity chromatography (IMAC) using HisTrap HP columns (GE Healthcare) coupled to an ÄKTAprime Plus protein purification system (GE Healthcare).

E. coli Tuner (DE3) cells (Invitrogen, Thermo Fisher Scientific) were transformed with anti-HR1 VL construction plasmids. VLs were growth in 4 L of lysogeny broth (LB) medium at 37 °C until OD_{600} of 0.4. Protein production was induced with 0.2 mM IPTG and lasted for 4 h at the same temperature. The recombinant proteins were expressed as inclusion bodies to yield a sufficient amount of protein for downstream analysis. Bacterial pellets were collected and resuspended in 160 mL of 50 mM HEPES pH 8.0, 1 M NaCl, 10 mM imidazole and 5 mM $CaCl_2$ previous to disruption through sonication during 30 min. The insoluble fraction was then resuspended in 80 mL of the same buffer and further disrupted by sonication during 20 min. The inclusion bodies were collected by centrifugation, resuspended in 40 mL of 50 mM HEPES pH 78.0, 1 M NaCl, 10 mM imidazole, 2 M urea and 5 mM $CaCl_2$ and subjected to sonication to remove soluble contaminants. Finally, insoluble proteins were resupended in 25 mL of 50 mM HEPES pH 8.0, 1 M NaCl, 10 mM imidazole, 6 M urea, 1 mM 2-mercaptoethanol, 5 mM $CaCl_2$ and solubilized overnight at 4 °C with gentle agitation. Soluble denatured proteins were recovered after centrifugation and purified by IMAC, using His GraviTrap columns (GE Healthcare) according to manufacturer's instructions.

Purified VLs underwent buffer exchanging (refolding in case of anti-HR1 VLs) to 20 mM HEPES pH 7.4, 0.1 M NaCl and 5% (v/v) glycerol using PD-10 Desalting Columns (GE Healthcare) according to manufacturer's instructions. Purified VLs were kept at -80 °C for long-term storage or at 4 °C for up to one month with no visible aggregation or loss of function. The protein yield was measured by Bradford method according to manufacturer's instructions. Protein purity was as assessed by SDS-PAGE.

VLs expression and purification were analyzed by Western blot as described,[350] where 20 µg of bacterial extract or 500 ng of purified protein were loaded into 12% Bis-Tris acrylamide gel.

2.3.6 Inhibition assays

HIV-1 laboratory-adapted strain NL4-3 (HIV-1_{NL4-3}) production was performed as described[354] and 50% tissue culture infectious dose ($TCID_{50}$) determined as Borrego *et al*.[343] HIV-2 and HIV-1 primary isolates from subtypes J and H were obtained from Borrego *et al*.[343] HIV-1 variant NL4-3 D36G (parental) and HIV-1 variants resistant to T-20 NL4-3 (D36G) V38A/N42D and V38A/N42T (NIH AIDS Reagent Program) were propagated accordingly to Borrego *et al*.[343] HIV-1 primary isolates from subtypes B and C were obtained from Calado *et al*.[355]

For all the inhibition assays, viruses or HeLa243*env*/HeLa273Δ*env* cells were incubated with titrated amounts of the inhibitors during 1 h at 37 ºC prior to infection and HIV infectivity was measured 48 h postinfection. In Jurkat E6-1 (NIH AIDS Reagent Program) inhibition assay, HIV $p24^{CA}$ concentrations were measured by ELISA (NCI Frederick, MD, USA) according to manufacturer's instructions. In inhibition assays with TZM-bl cells (NIH AIDS Reagent Program),[343] luciferase or β-galactosidase expression was quantified with the One-Glow luciferase assay substrate reagent (Promega, USA) according to manufacturer's instructions or as described in Da Silva *et al*,[43] respectively. Cell-cell fusion assay was adapted from Schwartz *et al*.[356] HeLa243*env* or HeLa273Δ*env* cells[356] were co-cultured at a 1:1 ration with MAGI cells (NIH AIDS Reagent Program) in the presence of inhibitors. After 48 h, β-galactosidase expression was quantified as described.[43] Peripheral blood mononuclear cells' (PBMCs) isolation, maintenance and inhibition assays were performed as previously described[355] with the following exception: at seven days postinfection, HIV $p24^{CA}$ concentrations were measured by ELISA (NCI Frederick, MD, USA). The 50% inhibitory concentration (IC_{50}) estimation and statistical analysis were performed as described.[164,343]

2.3.7 Liposome preparation

Large unilamellar vesicles (LUV) with controlled size and composition were prepared as described.[357] Formulations composed of 1-palmitoyl-2-oleyl-sn-glycero-3-phosphocholine (POPC; AvantiAvanti Polar Lipids; Alabaster, AL, USA) and a mixture of POPC and cholesterol (Chol; Sigma-Aldrich; St. Louis, MO, USA) in a 2:1 ratio were prepared. Each lipid formulation was first solubilized with spectroscopic grade

chloroform (Merck; Darmstadt, Germany) in a round bottom flask. A thin lipid film was formed after solvent evaporation under nitrogen flow and then in vacuum. Rehydration with 10 mM HEPES pH 7.4, 150 mM NaCl, 5% (v/v) glycerol and a series of 8-10 freeze/thaw cycles yielded a multilamellar vesicle (MLV) suspension. This suspension was then extruded through a 100 nm pore polycarbonate membrane (Whatman/GE Healthcare; Kent, United Kingdom) using a LiposoFast-Basic plus Stabilizer setup (Avestin; Mannheim, Germany). The resulting LUV suspension was used in fluorescence spectroscopy experiments.

2.3.8 Partition experiments

A 15 mM LUV stock solution was successively added to 10 μM of either F63 or parental VL for final concentrations ranging from 0 to 5 mM. The sample was incubated for 10 min between each addition. The extent of sdAb partition was followed by steady state fluorescence emission performed in a Varian Cary Eclipse spectrofluorometer (Palo Alto, CA, USA). A 280 nm radiation wavelength was used to excite tryptophan (Trp) and tyrosine (Tyr) residues. sdAb emission profiles were collected in the range between 300 and 450 nm. Fluorescence intensity was corrected for dilution, solvent background and light scattering.[358] Partition constants (K_p) were determined by fitting the emission spectra integral (I) with the partition equation.[359]

$$I = \frac{I_W + K_p \gamma_L [L] I_L}{1 + K_p \gamma_L [L]} \quad (1)$$

In this equation, I_W and I_L are the emission spectra integral of the sdAb in the aqueous and lipid phase, respectively; γ_L is the lipid molar volume and [L] is the lipid concentration.

2.3.9 Membrane dipole potential sensing

LUV were pre-incubated with di-8-ANEPPS (Sigma-Aldrich; St. Louis, MO, USA) overnight with agitation, to ensure efficient probe incorporation into the lipid membrane. The final probe to lipid ratio was 1/50 (2 mol% di-8-ANEPPS relative to lipid). Assays were performed with final concentrations of 200 μM for LUV, 4 μM for di-8-ANEPPS and 9 μM for each sdAb, assayed separately. Variations in the membrane dipole potential were monitored by di-8-ANEPPS excitation spectra deviations. Experiments were performed in a Varian Cary Eclipse spectrofluorometer. Spectra were collected between 380 and 580 nm, with a fixed emission wavelength at 670 nm to avoid membrane fluidity artifacts.[358] Excitation was corrected for solvent background noise.

Controls consisted in di-8-ANEPPS-stained LUV at the same final concentration but without the addition of the sdAb. These controls were used to obtain the differential excitation spectra. This was possible by subtraction of the normalized controls from normalized sample. The normalization was to the individual spectrum integral.

2.3.10 Affinity measurements

F63 binding kinetics was determined by surface plasmon resonance (SPR), using a BIAcore 2000 (BIAcore Inc.). Briefly, N36 was captured on a CM5 sensor chip (GE Healthcare) using amine coupling at ~1000 resonance units. Serial dilutions of F63 were injected for 4 min and allowed to dissociate for 10 min before matrix regeneration using 10 mM glycine pH 1.5. Signal from an injection passing over an uncoupled cell was subtracted from an immobilized cell signal to generate sensorgrams of the amount of F63 bound as a function of time. The running buffer was used for all sample dilutions. BIAcore kinetic evaluation software (version 3.1) was used to determine K_D from the association (K_a) and dissociation rates (K_d) using a one-to-one binding model. $VL_{parental}$ was used as negative control.

2.4 Results

2.4.1 Selection of anti-HIV VLs with elongated complementary-determining region 1 and complementary-determining region 3

In contrast to regular binding regions, long and flexible CDR3 of camelid heavy-chain antibodies[25] can successfully target hidden and non-standard (immune-evasion) epitopes.[31,344] We proposed to translate these CDR features to a non-camelid scaffold—a rabbit kappa VL domain [348] derived from a previously selected single-chain variable fragment (scFv).[349] In addition to the stability already attributed to VL sdAbs,[47,48,63,69] their solubility seems less affected by sequence variation in CDRs than VH domains.[47] We chose a non-human domain due to the extensive CDR3 length heterogeneity naturally present in the kappa light-chains of rabbit antibodies, in contrast with human ones.[360,361] Furthermore, this particular VL domain was shown to be highly stable and soluble in the absence of its counterpart VH domain.[348] The naturally longer and most exposed CDR of parental VL (CDR3; Fig. 2.1A), hereafter named $VL_{parental}$, was grafted with a series of long CDRs containing a well-characterized paratope for hen egg-white lysozyme (HEL)[30] flanked by sequences of serines/glycines. These small amino acids—major contributors

to conformational flexibility of antibody CDRs[79,362]—were added to elongate the original CDR3 of 11 amino acids (aa) to 22, 26 or 30 aa (Fig. 2.1B).

Evaluation of HEL targeting by each VL construction ($VLCDR3_{22}$, $VLCDR3_{26}$ and $VLCDR3_{30}$) demonstrated that only $VLCDR3_{22}$ showed binding to HEL antigen **(**HEL binding was 0.86 ±0.06 for $VLCDR3_{22}$ vs 0.07 ±0.03 for VL**;** $p<0.001$; Fig. 2.1C). Longer $CDR3_{26}$ and $CDR3_{30}$ may have result in non-functional conformations of CDR3 or the entire VL domain. The binding presented by the $VLCDR3_{22}$ was specific for HEL, not observed for BSA ($Abs_{405\ nm}$ was 0.86 ±0.06 for HEL binding vs 0.06 ±0.04 for BSA binding; $p<0.001$; Fig. 2.1C). Additionally, $VLCDR3_{22}$ showed a dose-dependent binding to HEL (EC_{50} was 159.5 ±1.1 nM; Fig. 2.1D). We also performed a competitive ELISA by incubating $VLCDR3_{22}$ with soluble HEL (competitor), prior to immobilized-HEL binding, to better characterize this VL construction specificity. As expected, $VLCDR3_{22}$ binding to immobilized-HEL decreased as the amount of HEL competitor increased (Fig. 2.1D). No HEL binding was observed for $VL_{parental}$. Besides $VLCDR3_{22}$ specificity towards HEL antigen, this result further proves VL capacity to bind soluble HEL in addition to immobilized-HEL. Finally, we analyzed $VLCDR3_{22}$ soluble expression in bacteria. Surprisingly, $VLCDR3_{22}$ showed higher soluble expression than $VL_{parental}$ for all tested post-induction time points (Fig. 2.1E). These results suggest that elongated-CDR sdAbs remain functional for antigen binding and could be more stable than standard VL domains.

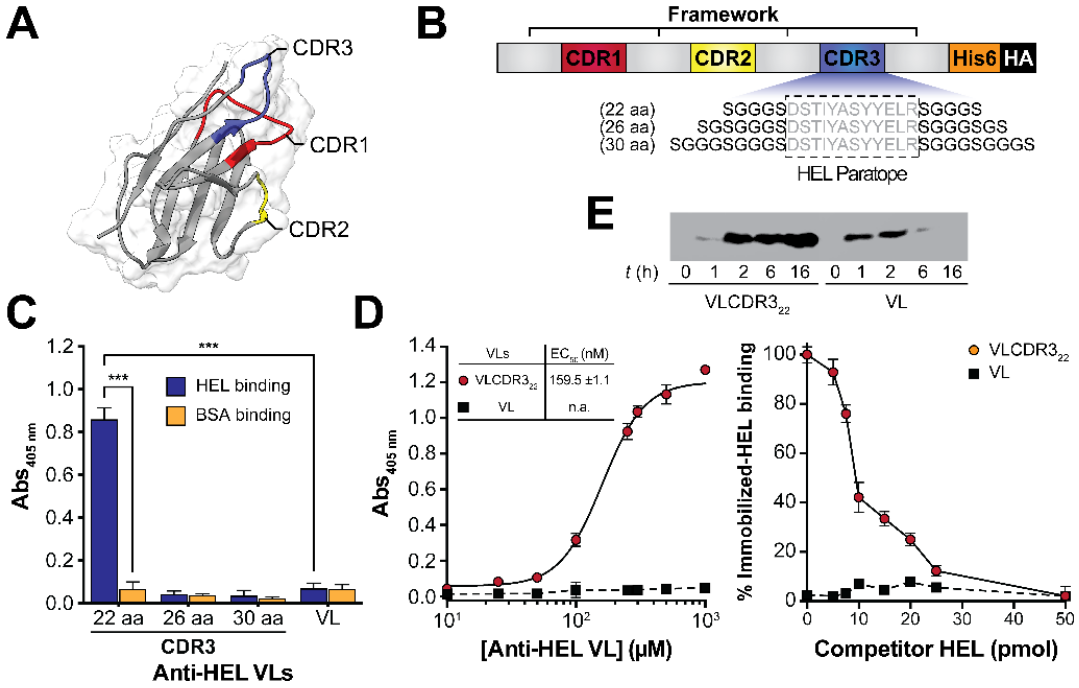
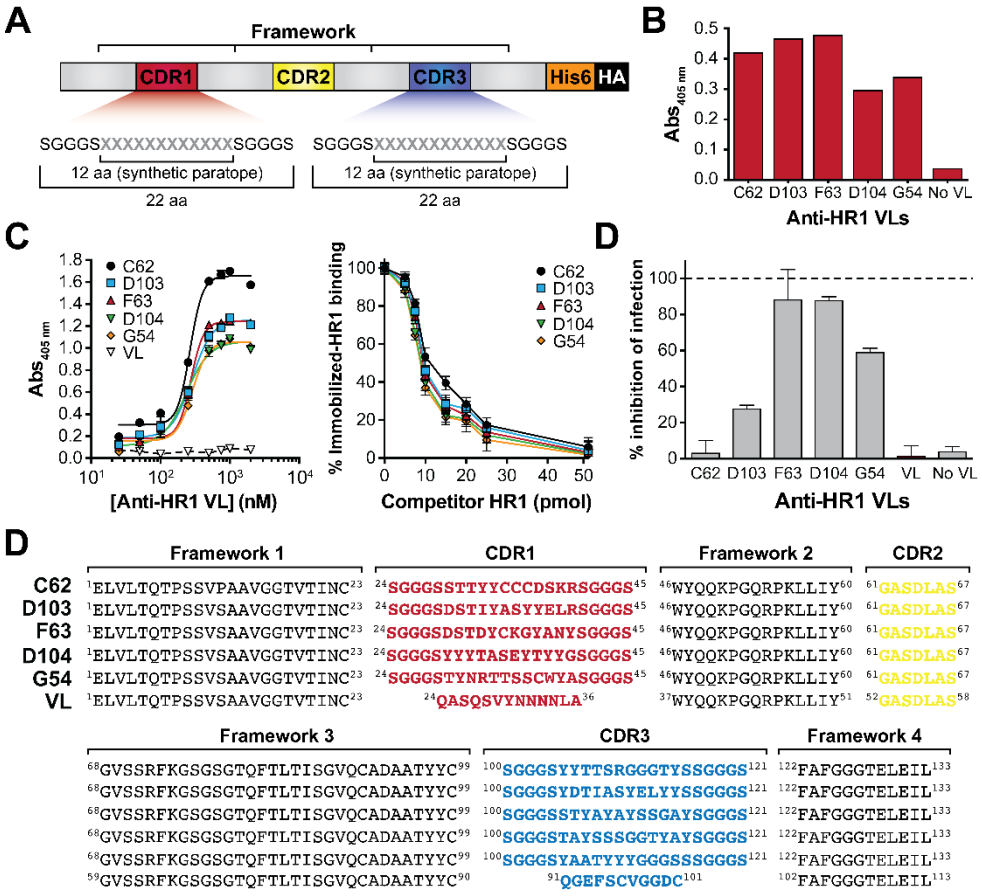


Figure 2.1: Functional analysis of VLs with elongated CDR3.
A) Three-dimensional structure prediction of $VL_{parental}$ by the I-TASSER protein structure homology-modelling server [11]. CDR1 is represented in red, CDR2 in yellow and CDR3 in blue. Framework is represented in grey. The molecular structure was represented using the UCSF chimera package [363]. **B)** Schematic representation of VLs specific for hen egg-white lysozyme (HEL) with CDR3 of different sizes. HEL paratope is represented in grey at the central region of CDR3. Different serines/glycines sequences were added to the flanks of HEL paratope to increase the flexibility and size of the CDR to 22, 26 or 30 aa. The hexa-histidine tail (His6) was used for further purification of the VLs and HA tag for detection. **C)** Selection of functional anti-HEL VLs. Binding and specificity of each VL construction with distinct CDR3 sizes were evaluated by ELISA using HEL and BSA as antigens and the $VL_{parental}$ (VL) as negative control. VL domains were incubated with HEL or BSA-coated wells. Nonspecific binders were washed out. Bound VL domains were detected by HRP conjugated anti-HA antibody. Data are displayed as Abs measurement at 405 nm. Error bars correspond to standard deviation (n = 3; * p<0.05; *** p<0.001; **** p<0.0001; t-test). **D)** Binding analysis of VLCDR322 to HEL. **(Left)** Increasing concentrations of VLCDR322 and $VL_{parental}$ (VL) were incubated with HEL or BSA-coated wells. Data are displayed as Abs measurement at 405 nm. To facilitate data representation, HEL binding was calculated according to the following formula: AbsHEL-coated well-AbsBSA-coated well. Error bars correspond to standard deviation (n = 3) and n.a. to not applicable. **(Right)** Competitive ELISA. VLCDR322 and $VL_{parental}$ (VL) were pre-incubated with increasing quantities of soluble HEL (competitor HEL) at 37 ºC. After 1 h, this mixture was incubated with HEL-coated wells (immobilized-HEL). Data are displayed as percentage of immobilized-HEL binding (no competitor/immobilized-HEL = 100%) according to formula: $[(Abs_{competitor/immobilized\text{-}HEL}-Abs_{competitor/immobilized\text{-}BSA})/(Abs_{nocompetitor/immobilized\text{-}HEL}-Abs_{nocompetitor/immobilized\text{-}BSA})]*100$. Error bars correspond to standard deviation (n = 3). **E)** Soluble expression of VLCDR322 in bacteria. The soluble expression yields of VLCDR322 and $VL_{parental}$ (VL) were compared by Western-Blot using the HRP conjugated anti-HA antibody for detection at five time points post-induction (t).

After validation of $VL_{parental}$ functionality in the presence of an elongated CDR3, we used it as a scaffold for synthetic library construction. This library was designed to select anti-HIV minimal antibody fragments with high-affinity towards a cryptic HR1 region. We chose to elongate both CDR1 and CDR3 to increase theoretically the affinity of the selected sdAbs and at the same time avoid unspecific binding from the original CDR1. For the CDRs library construction, we used the previously validated strategy for HEL paratope grafting (Fig. 2.2A), restricted randomization of the central 12 aa with a degenerate codon (DVN) that only encodes for 12 of the canonical 20 aa. As most

encoded amino acids by DVN codon were described as abundant in natural CDRs and antigenic contacts [83], we expected to improve the selection of high-affinity sdAbs. Moreover, these amino acids seem to be sufficient to generate high-affinity and specific minimalist synthetic binders.[83,362] A library of ~8.0x10^9 clones was generated, cloned and selected by phage display against a crucial-to-fusion, difficult-to-access and well conserved sequence on HR1 (HR1$_{546–581(HXB2)}$), named N36[116] (Fig. 2.8 Supplementary Information). Apart from the cryptic nature of the entire HR1 region, N36 comprises residues of a particularly deep cavity, named hydrophobic pocket, described as highly conserved and a hot-spot for neutralization of HIV-1 infection [116,364]. Despite the therapeutic interest, this pocket is particularly difficult to target in an infection context due to its extreme concave conformation. As shown in Fig. 2.2B, we isolated five VLs with strong binding to HR1 from the 329 clones screened by ELISA. A further characterization of HR1 binding showed a dose-dependent binding for all five selected VLs (Fig. 2.2C and Table 2.2 Supplementary Information). A competitive ELISA demonstrated that the five VLs showed a decreased binding to immobilized-HR1 as the soluble HR1 amount increased, confirming VLs specificity of recognition (Fig. 2.2C).



***F*igure 2.2: Selection of anti-HIV VLs from the constructed synthetic library.**

A) Schematic representation of VLs synthetic library. CDR1 and CDR3 were randomized in the central 12 amino acids (aa) represented by the X letter in grey. Serines/glycines sequence was added to the flanks to provide flexibility. The hexa-histidine tail (His6) was used for further purification of the VLs and HA tag for detection. **B)** Anti-HR1 VLs selection. The anti-HR1 VLs were expressed in bacteria and cell extracts incubated with N36 region of HR1 or BSA in ELISA plates. No VL represents cell extract with no VL expression. Nonspecific VLs were washed out. Bound VLs were detected by HRP-conjugated anti-HA antibody. The five phage-selected VLs out of 329 that presented highest binding values to HR1 are represented. Data are displayed as Abs measurement at 405 nm. To facilitate data representation, HR1 binding was calculated according to the following formula: $Abs_{HR1\text{-}coated\ well}$-$Abs_{BSA\text{-}coated\ well}$. **C)** HR1 binding analysis. **(Left)** Increasing concentrations of the purified five selected VLs and control $VL_{parental}$ (VL) were incubated with HR1 or BSA-coated wells. Data are displayed as Abs measurement at 405 nm. To facilitate data representation, HR1 binding was calculated according to the following formula: $Abs_{HR1\text{-}coated\ well}$-$Abs_{BSA\text{-}coated\ well}$. Error bars correspond to standard deviation (n = 3). **(Right)** Competitive ELISA. The five selected VLs and $VL_{parental}$ (VL) were pre-incubated with increasing quantities of soluble N36 region of HR1 at 37 ºC. After 1 h, this mixture was incubated with N36-coated wells (immobilized-HR1). Data are displayed as percentage of immobilized-HR1 binding (no competitor/immobilized-HR1 = 100%) according to formula: $[(Abs_{competitor/immobilized\text{-}HR1}\text{-}Abs_{competitor/immobilized\text{-}BSA})/(Abs_{nocompetitor/immobilized\text{-}HR1}\text{-}Abs_{nocompetitor/immobilized\text{-}BSA})]^{*}100$. Error bars correspond to standard deviation (n = 3). **D)** Selection of antiviral VLs. TZM-bl cells were infected with HIV-1NL4-3 laboratory-adapted strain in the anti-HR1 VLs presence (C62, D103, F63, D104 and G54), after 3 h of inhibitors-viruses incubation. TZM-bl cell line expresses β-galactosidase and luciferase genes under control of HIV-1 promoter (LTR)—activated in the presence of HIV Tat protein (infection). $VL_{parental}$ (VL) represents the negative control. No VL represents cell extract with no VL expression. Data are displayed as percentage of infectivity inhibition (virus/no inhibitors = 0% inhibition; no virus/no inhibitors = background) according to the formula: $[(Abs_{virus/inhibitors}\text{-}Abs_{background})/(Abs_{virus/noinhibitors}\text{-}Abs_{background})]^{*}100$. Error bars correspond to standard deviation (n = 2). E) Amino acid sequences of the five anti-HR1 VLs: C62, D103, F63, D104, G54 and control $VL_{parental}$ (VL). CDR1 is highlighted in red, CDR2 in yellow and CDR3 in blue. VLs backbone composed of four frameworks is represented in grey.

To assess the antiviral activity of selected VLs, we performed a second screening ("functional screening") against HIV-$1_{NL4\text{-}3}$—encoding the G547D mutation responsible for less susceptibility to T-20 fusion inhibitor [162]. From the five anti-HR1 VLs, F63 and D104 inhibited HIV-$1_{NL4\text{-}3}$ infectivity by ~90% (Fig. 2.2D) and were selected for further characterization of antiviral activity. Except for CDR1 of VL D103 that was not randomized, DNA sequencing analysis confirmed that all CDR1 and CDR3 sequences of the five anti-HR1 VLs were unique (Fig. 2.2E). F63 and D104 VLs were expressed and purified in high yield and used for further functional characterization (Fig. 2.9 Supplementary Information).

2.4.2 Epitope mapping of antiviral VLs

Epitopes of F63 and D104 were mapped by ELISA, using a set of 10 overlapping synthetic peptides covering the template HR1, and compared to the T-20 binding region (Fig. 2.3A). Both VLs exhibited similar target sequences within the central region of N36 with short overlap at the C-terminus of T-20 binding region (Fig. 2.3B, C). F63 showed the strongest binding to peptide 5 (NH_2-EAQQHMLQLTVWGIK-COOH), suggesting that it might contain its epitope (Fig. 2.3B). D104 showed similar binding to peptides 5 and 6 (Fig. 2.3B), indicating that its epitope is located within the overlapping sequence of 11 aa NH_2-HMLQLTVWGIK-COOH (Fig. 2.3A).

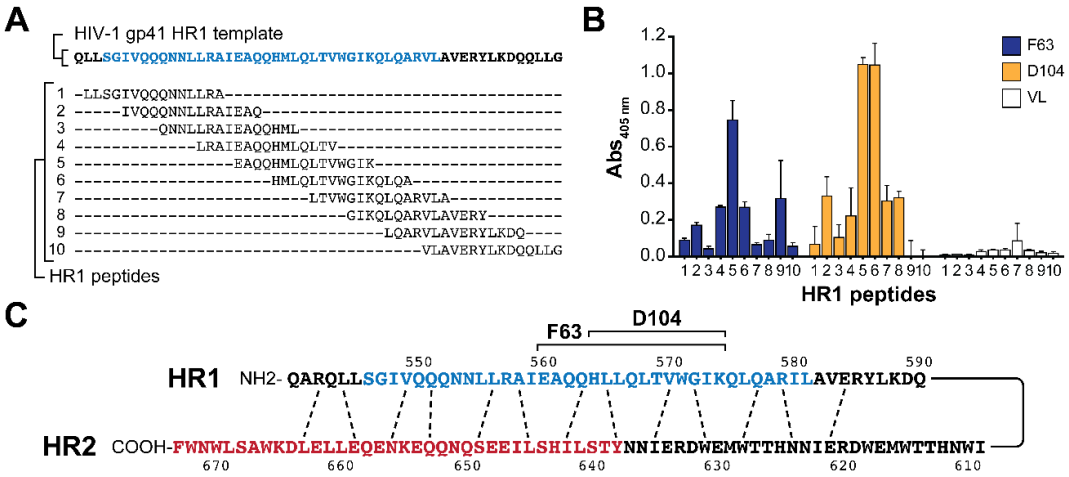


Figure 2.3: Epitope mapping of antiviral VLs.
A) Amino acid sequences of the 10 peptides (15-mer) representing the HR1 template (N36 in blue) used as antigens to map the F63 and D104 epitopes. Each peptide comprises 15 residues, 11 aa overlapping the subsequent peptide and an overhang of 4 aa at N-terminal region. **B)** Epitope mapping of F63 and D104 by ELISA, using 10 overlapping peptides of HR1 region and BSA as antigens and $VL_{parental}$ (VL) as negative control. Antiviral VLs were incubated with HR1 peptides or BSA-coated wells. Only in the wells coated with epitope-containing HR1 peptides, VLs were detected with anti-HA-HRP antibody. In the remaining wells, VLs were washed out. Data are displayed as Abs measurement at 405 nm. To facilitate data representation, HR1 binding was calculated according to the following formula: AbsHR1-coated well-AbsBSA-coated well. Error bars correspond to standard deviation (n = 3). **C)** Location of predicted F63 and D104 epitopes in HR1 region. Amino acid residues in red constitute the T-20 origin and sequence. Amino acid residues highlighted in blue represent the N36 region. Dash lines represent HR1-HR2 interactions.

2.4.3 Antiviral activity of VLs

The antiviral activity of VLs was first compared with T-20 peptide against the HIV-1_{NL4-3}. T-20 also binds HR1 impairing the virus-cell fusion driven by HR1-HR2 interaction, an inhibition mechanism we reasoned to be similar to selected anti-HIV VLs. As shown in Fig. 2.4A, both F63 and D104 strongly inhibited HIV-1_{NL4-3} infection in TZM-bl cells (IC_{50} was 0.5 ±0.2 nM for F63 and 9.7 ±5.4 nM for D104). Remarkably, F63 was more active against HIV-1_{NL4-3} than D104 and T-20 (IC_{50} was 0.5 ±0.2 nM for F63 vs 9.7 ±5.4 nM for D104 and 3.1 ±1.9 nM for T-20). As expected, the $VL_{parental}$ did not inhibit HIV-1_{NL4-3}. F63 and D104 also strongly inhibited HIV-1 infection similarly to T-20 in Jurkat cells (IC_{50} was 0.1 ±0.01 nM for F63, 0.6 ±0.1 nM for D104 and 0.1 ±0.01 nM for T-20; Fig. 2.4B). No cytotoxicity was observed when either TZM-bl or Jurkat cells were incubated with the highest concentration of the VLs (Fig. 2.10 Supplementary Information). We also assessed F63 antiviral activity as a dimer. Surprisingly, F63 dimer did not inhibit HIV-1_{NL4-3} infection (data not shown), which is probably related to steric restrictions in F63 epitope access.

We then asked whether F63 and D104 could inhibit cell-cell fusion between Env-positive cells (HeLa243*env*) and adjacent CD4-expressing cells (MAGI) [356]. Similarly to T-20 and in contrast to $VL_{parental}$, F63 and D104 impaired HeLa cell-cell fusion in a

concentration-dependent manner (Fig. 2.4C). Fusion between control HeLa273Δ*env*—without Env expression—and MAGI cells did not occur (data not shown). These results emphasize HIV-1 fusion as the target of F63 and D104.

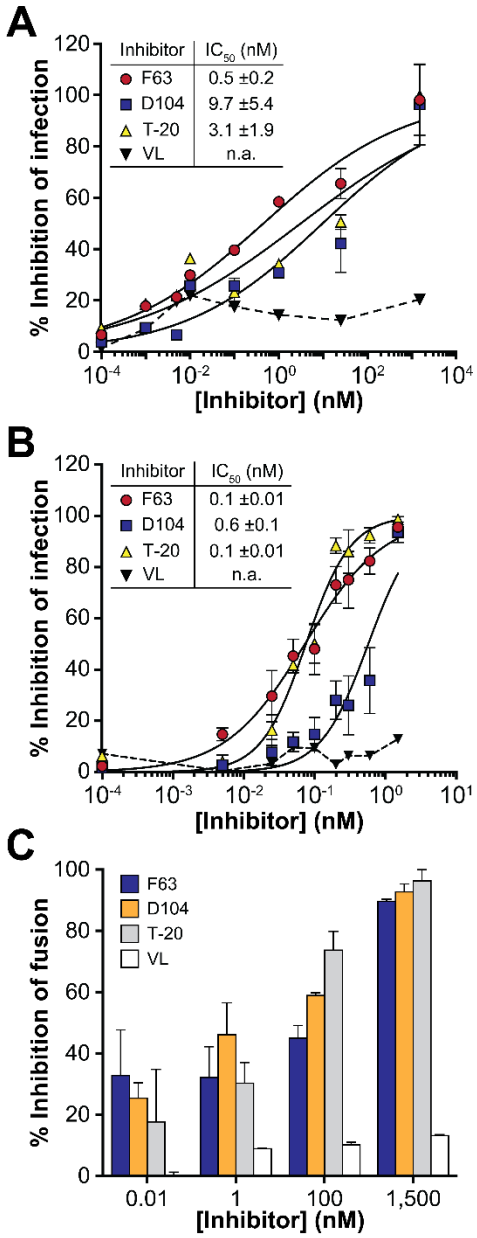


Figure 2.4: Antiviral activity of VLs.
Percentage of viral infection inhibition was assessed against the laboratory-adapted strain HIV-1NL4-3 in TZM-bl **(A)** and Jurkat cells **(B)**. Titrated amounts of VLs were incubated with HIV-1NL4-3 for 1 h at 37 ºC prior to infection. After 48 h, HIV infectivity was evaluated by β-galactosidase activity measurement (TZM-bl) or p24CA quantification (Jurkat). Data are displayed as percentage of infectivity inhibition (virus/no inhibitors = 0% inhibition; no virus/no inhibitors = background) according to the formula: $[1-(Abs_{virus/inhibitors}-Abs_{background})/(Abs_{virus/noinhibitors}-Abs_{sbackground})]*100$. Error bars correspond to standard deviation (n = 3). **C)** Cell-cell fusion assay. HeLa cells presenting functional gp120/gp41 complexes at cell surface and expression of HIV-1 Tat protein (HeLa243env) were co-cultured during 48 h with CD4-expressing HeLa cells (MAGI) in the inhibitors presence. Prior to co-culturing, HeLa243env were incubated with titrated amounts of the VLs for 1 h at 37 ºC. MAGI cells have an integrated β-galactosidase gene under control of LTR promoter, similarly to TZM-bl cells. Fusion inhibition was assessed by β-galactosidase activity measurement. Data are displayed as percentage of fusion inhibition (HeLa243env cells/no inhibitors = 0%; no HeLa243env cells/no inhibitors = background) according to the formula: $[1-(Abs_{HeLa243env/inhibitors}-Abs_{background})/(Abs_{HeLa243env/noinhibitors}-Abs_{background})]*100$. Error bars correspond to standard deviation (n = 3).

To test the hypothesis that F63 and D104 were as active as T-20 towards clinically relevant HIV isolates, the IC_{50} of VLs was evaluated against two HIV-1 and HIV-2 primary isolates in TZM-bl cells. HIV-1 primary isolates belong to distinct subtypes of the major HIV-1 group M, clade J (93AOHDC250) [164] and clade H (93AOCA251) [164]. HIV-2 primary isolates 03PTHCC12 and 10PTHSMNC [164] belong to the most prevalent HIV-2 group [365] (Group A, ~90% worldwide). Despite the divergent HR1 sequences of the HIV-1 primary isolates (Fig. 2.8 Supplementary Information), F63 neutralized both viruses (IC_{50} was 402 ±46 nM for 93AOHDC250 isolate and 469 ±41 nM for 93AOCA251 isolate; Fig. 2.5A). Although F63 did not inhibit these HIV-1 primary isolates as potently as T-20, IC_{50} values are in the nanomolar range for both HIV-1 subtypes (IC_{50} was 402 ±46 nM for F63 vs 1.3 ±0.4 nM for T-20 for 93AOHDC250 isolate; 469 ±41 nM for F63 vs 0.4 ±0.1 nM for T-20 for 93AOCA251 isolate; Fig. 2.5A and Fig. 2.5B). This fact still supports F63 as a potent inhibitor of these two HIV-1 primary isolates. In contrast to T-20, F63 also inhibited the two HIV-2 primary isolates in the nanomolar range (IC_{50} was 460 ±19 nM for F63 vs 2,855 ±483 nM for T-20 for 03PTHCC12 isolate; IC_{50} was 433 ±38 nM for F63 vs 266 ±25 nM for T-20 for 10PTHSMNC isolate; Fig. 2.5C and Fig. 2.5D). In contrast to F63, D104 did not inhibit either HIV-1 or HIV-2 primary isolates (data not shown). These results suggest that in addition to inhibition of HIV-1 non-B subtypes, F63 can be a potent inhibitor of HIV-2 isolates. We also tested the antiviral potency of F63 in PBMCs against a panel of HIV-1 primary isolates from the most prevalent subtypes B (developed countries) and C (developing countries). As shown in Table 2.3 Supplementary Information, the IC_{50} values obtained for F63 were similar to T-20 in the nano-picomolar range. These results indicate that F63 also potently inhibits isolates from the most prevalent HIV-1 subtypes in primary lymphocytes, with comparable activity to T-20.

Overall, F63 was as active as T-20 as judged by the IC_{50} against all tested HIV isolates (no significant *p*-value; Fig. 2.5E). Moreover, F63 was not significantly less active than T-20 either for HIV-1 or HIV-2 primary isolates (no significant *p*-value; data not show). These results indicate that F63 presents an antiviral activity similar to T-20.

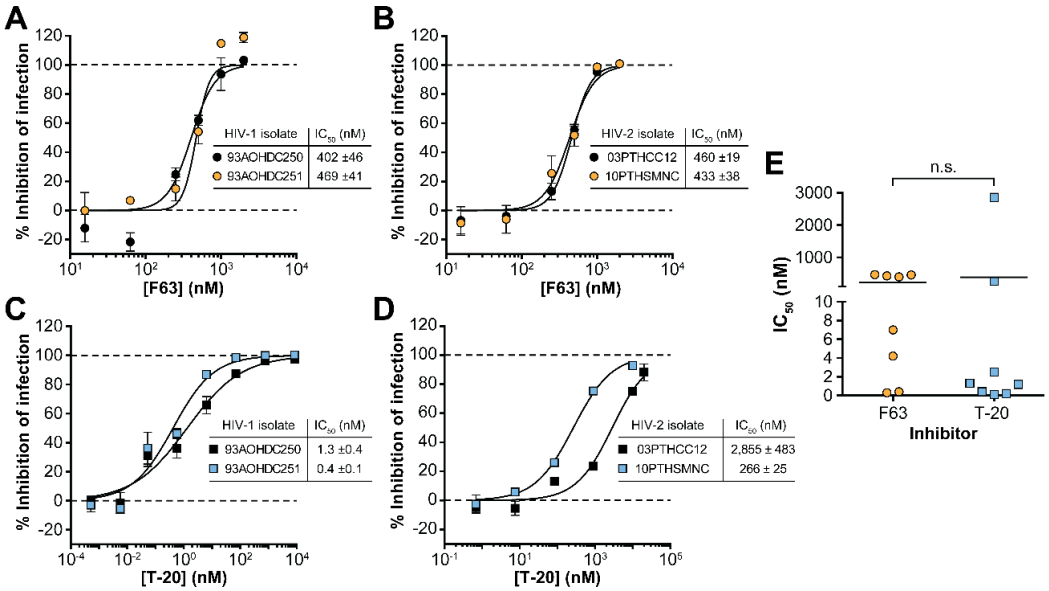


Figure 2.5: Antiviral activity of VLs against HIV primary isolates.
Antiviral activity of F63 was compared to the FDA-approved T-20 against HIV-1 **(A and B)** and HIV-2 **(C and D)** primary isolates. Titrated amounts of VLs were incubated with HIV primary isolates for 1 h at 37 ºC prior to TZM-bl cells infection. After 48 h, HIV infectivity was evaluated by luciferase activity measurement. Data are displayed as percentage of infectivity inhibition (virus/no inhibitors = 0%; no virus/no inhibitors = background) according to the formula: $[1-(LU_{virus/inhibitors}-LU_{background})/(LU_{virus/noinhibitors}-LU_{background})]*100$. Error bars correspond to standard deviation (n = 4). **E)** Mean IC50 values of F63 and T-20 for all tested HIV primary isolates ($p<0.05$; *** $p<0.001$; **** $p<0.0001$; n.s. = no significant; t-test). Bars represent median values.

To test F63 neutralizing activity against HIV-1 strains resistant to T-20, we evaluated the susceptibility of two HIV-1 variants displaying well-defined mutations for T-20 resistance [162,366]. HIV-1 variants resistant to T-20 derived from HIV-1 NL4-3 D36G (parental) susceptible to T-20 [162,366]. F63 presented no fold-increase of IC_{50} for all tested HIV-1 variants resistant to T-20 relative to parental HIV-1 (IC_{50} fold-increase was 0.66 for NL4-3 (D36G) V38A/N42D; no significant *p*-value and 0.70 for NL4-3 (D36G) V38A/N42T; no significant *p*-value; Fig. 2.6). In contrast, T-20 IC_{50} was reported to present a fold-increase of approximately $3.94x10^3$ and $1.61x10^4$ for HIV-1 variants NL4-3 (D36G) V38A/N42D and V38A/N42T, respectively, comparing to the parental HIV-1 [343]. These data suggest that F63 could constitute an alternative in the treatment of patients infected with HIV-1 strains resistant to T-20.

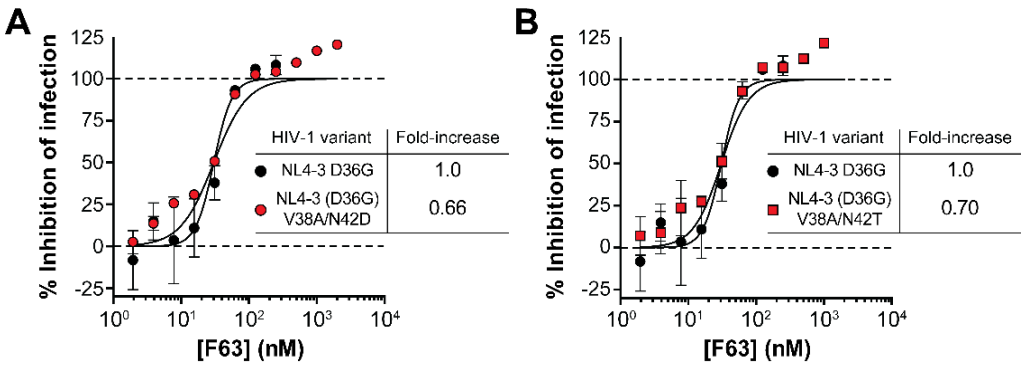


Figure 2.6: Antiviral activity of F63 against HIV-1 strains resistant to T-20.
Antiviral activity of F63 against HIV-1 variants resistant to T-20 NL4-3 (D36G) V38A/N42D **(A)** and V38A/N42T **(B)** was compared to HIV-1 variant susceptible to T-20 D36G (parental). Titrated amounts of F63 were incubated with HIV-1 variants for 1 h at 37 ºC prior to TZM-bl cells infection. After 48 h, HIV infectivity was evaluated by luciferase activity measurement. Data are displayed as percentage of infectivity inhibition (virus/no inhibitors = 0%; no virus/no inhibitors = background) according to the formula: $[1-(LU_{virus/inhibitors}-LU_{background})/(LU_{virus/noinhibitors}-LU_{background})]*100$. Error bars correspond to standard deviation (n = 2). Fold-increase represents fold-increase of IC50 relative to NL4-3 D36G (parental).

2.4.4 F63 interaction with lipid membranes

Since T-20 antiviral mechanism is associated with membrane interaction, we also evaluated the lipid binding capacity of F63 against membrane model systems mimicking the major lipids of cellular membrane and cholesterol-rich viral envelope [367]. In the presence of lipid membranes, variations in the fluorescent residue emission are typically associated with protein-membrane interactions [368]. Taking advantage of the tryptophan residue (Trp; position 46 in F63 and 37 in $VL_{parental}$ Fig. 2.2E), we performed partition experiments based on the VL quantum yield variations. Fluorescence emission from F63 Trp decreased with increasing lipid concentrations of the both membrane models tested (Fig. 2.7A, B). The K_p values retrieved from data fitting with the partition formalism were in the same order of magnitude for viral envelope and cell membrane models (Table 2.4 Supplementary Information). The K_p correlates with the extent of protein interaction with the lipid membrane models—ratio between the concentration of a given molecule in two separate and immiscible phases. In contrast, we did not observe variations in the fluorescence emission of $VL_{parental}$ Trp (Fig. 2.7A, B). These results suggest that F63 interact with lipid membranes, independently of the cholesterol content. Spectral properties of lipophilic probes such as di-8-ANEPPS, which are responsive to variations in membrane dipole potential, can also be exploited to study protein-membrane interactions [359]. Excitation spectra of di-8-ANEPPS inserted in both membrane models underwent a red-shift to higher wavelengths—indication of a membrane dipole potential perturbation—only in the F63 presence (Fig. 2.7C, D). These observations complement the previous partition results and support the hypothesis that F63 has unique membrane-

interacting properties, unlike $VL_{parental}$. F63 also presented a binding affinity (K_D) of ~8 nM to N36 as determined by surface plasmon resonance (Table 2.5 Supplementary Information), establishing this VL as a high-affinity binder in the low-nanomolar range.

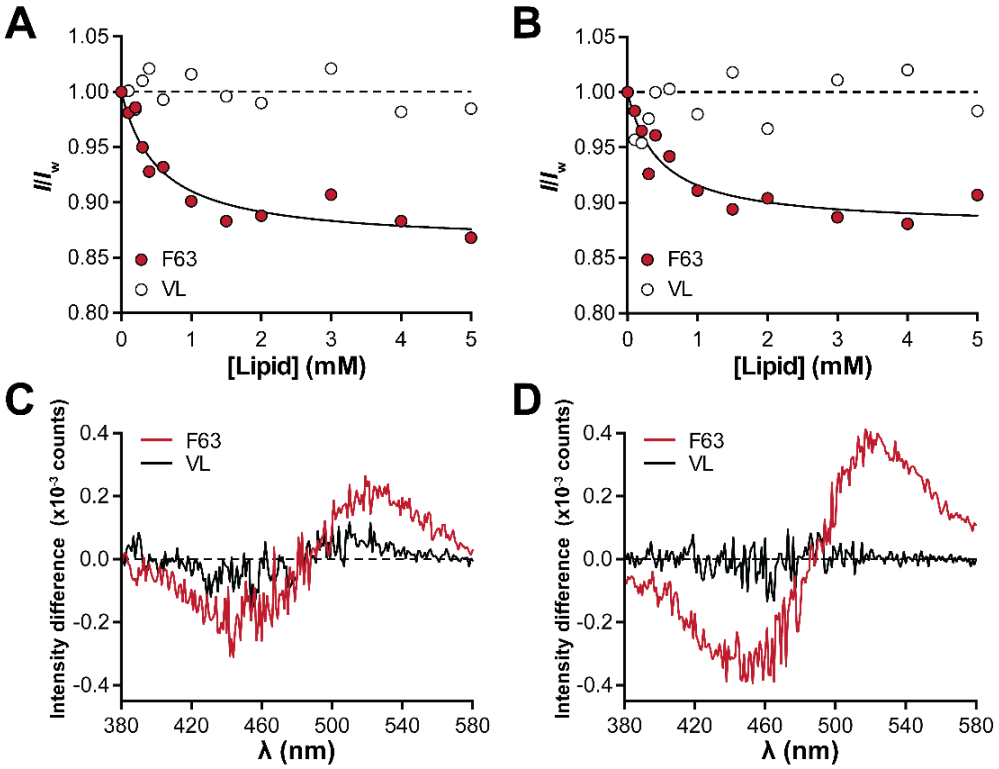


Figure 2.7: F63 membrane interactions.
Partition profiles of F63 and control VLparental (VL) towards 1-palmitoyl-2-oleyl-sn-glycero-3-phosphocholine (POPC; cellular membrane model) **(A)** and POPC:Cholesterol (2:1; virus envelope model) **(B)**. 10 μM of F63 and $VL_{parental}$ were titrated with small volumes of LUV up to final lipid concentrations, [L]. sdAb intrinsic fluorescence emission, I, was collected for each [L], and normalized to the respective emission in the aqueous media, IW. The line represents the best fit of equation 1 to one of three independent replicates. Differential excitation spectra of di-8-ANEPPS-labelled POPC **(C)** and POPC:Cholesterol (2:1) **(D)** liposomal membrane models in the presence of 9 μM of either F63 or control $VL_{parental}$ (VL). The respective amplitude is correlated with di-8-ANEPPS spectral shifts in response to membrane dipole potential perturbations. Graphs were obtained by subtraction of the normalized di-8-ANEPPS excitation spectra controls from the spectra in the presence of each VL (normalization to the respective spectrum integral). The presented spectra constitute one of three independent replicates.

2.5 Discussion

HIV entry inhibition is a key component of any antiviral therapeutic scheme leading to impairment of *de novo* infection. In this chapter, we selected a broad and potent HIV fusion inhibitor from a synthetic repertoire of VL domains. Several studies have shown that VLs tend to aggregate less [47,48,63] and present higher antigen-binding properties [67,70] than VH domains. In this study, we went further relative to others [68,369] and successfully selected a high-affinity VL with elongated CDRs. Our data suggest that the screening of

libraries containing CDRs-elongated antibody formats result in the selection of cryptic epitope binders, mimicking the longer and more flexible CDRs found in camelids [370].

Here, we took advantage of VLs reduced size to target a sterically restricted region on HR1 of gp41 (N36). Together with the corresponding HR2 region, N36 is sufficient to form the 6HB structure responsible for the HIV fusion [116]. Accordingly, a VL-N36 interaction would prevent the 6HB assembly, leading to HIV entry impairment. Epitope mapping of the most potent HIV inhibitors revealed two similar sequences within the N36 central region, previously described as part of a highly conserved cavity (hydrophobic pocket) essential for HR2 binding [116,120]. Despite the location of D104 target region within the F63 target sequence, antiviral activity against HIV primary isolates was only observed with F63. Thus, our data seem to indicate that targeting of D104 epitope is insufficient for broad neutralization of HIV. Nevertheless, since D104 affinity was not determined we cannot exclude that it may influence viral inhibition. On the other hand, F63 epitope represents a promising target with ~60% conservation amongst HIV-1 subtypes and even HIV types (Fig. S2). This predicted epitope is also distinct from the T-20 binding region that has a low genetic barrier to drug resistance, mainly the Gly-Ile-Val sequence ($HR1_{36-38(HXB2)}$) [162,371], as reinforced by the observed F63 inhibition of T-20 resistant HIV-1 strains. Moreover, a substitution of a single residue on ~70% of the F63 predicted epitope would lead to impaired or non-functional HIV-1 entry mutants as reported by *Sen* et al. [372] (Fig. S2). F63 epitope conservation and importance for HIV fusion together with the fact that this VL domain inhibited HIV-1 primary isolates from distinct subtypes similarly to T-20 and HIV-2 primary isolates highlight F63 potency and predict a high breadth for this inhibitor. Moreover, since T-20 has a limited activity on HIV-2 [165,343] F63 could constitute an alternative to the treatment of this HIV type.

The close proximity of gp41 to viral envelope and cellular membrane during HIV entry questions the role of membranes in gp41-targeting inhibitors mechanism. For example, T-20 shows considerable interaction with lipid membranes [346]. Also, broadly neutralizing antibodies 2F5 and 4E10 are capable of stable epitope binding through cross-reactive lipid interaction [373]. We have assessed F63 membrane interactions through fluorescence spectroscopy methodologies and identified its partition towards lipid membrane models. Interestingly, the $VL_{parental}$ was unable to interact with these models, suggesting that this property was acquired during CDRs randomization and is associated to CDR1 and/or CDR3. From a pharmacological standpoint, membranes interaction is a

desirable property of an inhibitor mechanism [374], enabling the establishment of local and transient reservoirs both in the viral envelope and cellular membrane. Furthermore, the F63 lack of a Fc immune-triggering domain avoids cross-reactivity with lipids, a significant drawback in 2F5 and 4E10 application [375].

To our knowledge, this is the first study presenting a synthetic VL sdAb designed as a potent inhibitor of HIV infection. Other fusion inhibitors with an antiviral activity similar to F63 were already described [168,376]. However, F63 combine the reduced molecular weight of small non-antibody inhibitors with the specificity and high-affinity of antibody paratopes and the excellent biophysical properties and versatility of antibody formats. Despite VHH antibody fragments were also identified as anti-HIV inhibitors [119,201,377–379], these variable domains target gp120 and were not synthetically randomized, being derived from llama immunization. Due to protease resistance and simplicity of sdAbs, F63 may also overcome major T-20 weaknesses such as oral administration preclusion, high production cost and short half-life [20,380]. Additionally, F63 potency and biodistribution may be further improved by several strategies such as coupling of effector molecules (enzymes and cytotoxic drugs) and inhibitor-targeting to the cholesterol-rich areas where HIV preferentially enters [381], for example through attachment of cholesterol-binding peptides.

To address the expected immunogenicity of a rabbit VL, we successfully humanized F63 by removing residues potentially recognized as T-cell epitopes (de-immunization) as described in Jones et al. [382]. Humanized F63 neutralized HIV-1_{NL4-3} laboratory-adapted strain similarly to rabbit F63 and proved to be more stable than its rabbit homolog due to alanine-substitution of unpaired cysteines performed along with F63 humanization (data not shown). It is conceivable that the rabbit-conserved cysteine at position 91—that forms an unusual disulfide bridge between variable and constant domains [383]—together with the cysteine selected in CDR1 sequence were major contributors to the insolubility of F63 during the purification process. Therefore, our library design strategy could benefit from the replacement of DVN by the NDT degenerate codon, which encodes fewer cysteine residues and does not yield stop codons.

In conclusion, we successfully developed a potent and broad fusion inhibitor of HIV-1 and HIV-2 infection using a VL sdAb as scaffold. We validated the selection of potent inhibitors based on a rational engineering strategy for synthetic library design. Our

findings also encourage exploration of CDRs elongation for the design or improvement of next-generation HIV inhibitors.

2.6 Acknowledgments

We thank C. Barbas III for kindly providing pComb3x plasmid, O. Schwartz for kindly providing the HeLa243*env* and HeLa273Δ*env* cells and Technophage for kindly providing modified pT7-FLAG-2 and purified $VL_{parental}$.

C.C-S., T.F., P.B., S.O., C.R., A.C., C.C., C.F., F.A-S., A.V., M.C. and J.G. conceived and designed the experiments. C.C-S., T.F., P.B., C.R., Q.S-C. and F.A-S. performed the experiments. T.F. performed the F63 interaction with membranes experiments. P.B. and C.R. performed the inhibition assays with the HIV primary isolates. P.B. performed the inhibition assays with HIV-1 resistant strains to T-20. Q.S-C. performed the inhibition assays on primary cells. F.A-S. performed the affinity measurements. C.C-S., T.F., P.B. and F.A-S analysed the data. C. C-S. and T.F. wrote the article.

This work was supported by grants HIVERA/0002/2013, PTDC/SAU-EPI/122400/2010 and VIH/SAU/0029/2011 from Fundação para a Ciência e a Tecnologia—Ministério da Educação e Ciência (FCT-MEC), Portugal. C.C-S. and T.F. were supported by FCT-MEC PhD fellowships SFRH/BD/73838/2010 and SFRH/52383/2013. F.A-S. and A.S.V. were supported by FCT Investigator Programme IF/01010/2013 and IF/00803/2012.

The following reagents were obtained through the NIH AIDS Reagent Program (Division of AIDS, NIAID, NIH): HIV-1 Subtype B (MN) Env Peptide Set; T-20, Fusion Inhibitor from Roche; pNL4-3 from M. Martin [366]; HIV-1 NL4-3 gp41 D36G Virus from Trimeris, Inc. [162,366]; HIV-1 NL4-3 gp41 (36G) V38A, N42D Virus from Trimeris, Inc. [162,366]; HIV-1 NL4-3 gp41 (36G) V38A, N42T Virus from Trimeris, Inc. [162,366]; TZM-bl from J. C. Kappes, X. Wu and Tranzyme Inc. [163,384–387]; HeLa-CD4-LTR-β-gal from M. Emerman [388] and Jurkat Clone E6-1 from A. Weiss [389].

2.7 Accession codes

The VL F63 sequence reported here has been deposited in the GenBank data base (accession number KT119563).

2.8 Supplementary Information

2.8.1 Supplementary Figures

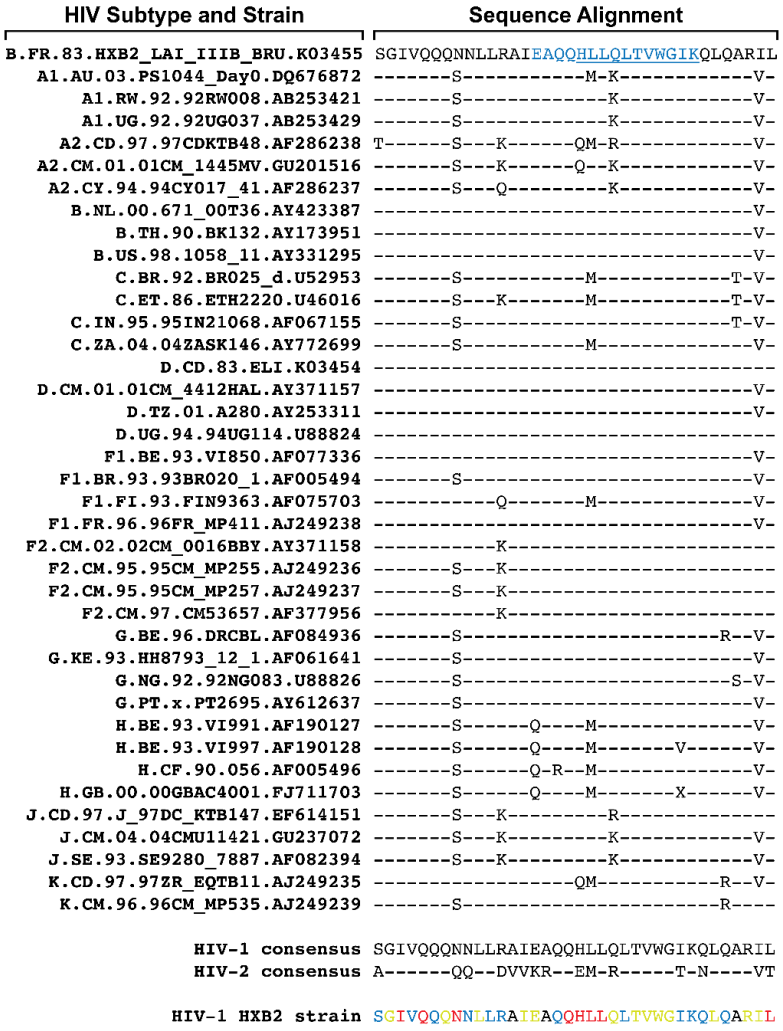


Figure 2.8: HR1 peptide conservation across HIV-1 subtypes and HIV types and N36 importance for HIV-1 entry.
Amino acid sequence alignment of N36 from reference strains of the diverse HIV-1 subtypes **(Top)**. Data based on 2014 edition of the HIV Sequence Database (**http://hiv-web.lanl.gov**). Dashes indicate sequence identity to HIV-1 consensus sequence of HXB2 strain from subtype B (first sequence). N36 sequence highlighted in blue corresponds to predicted F63 epitope. N36 underlined is D104 predicted epitope. Amino acid sequence alignment of N36 from HIV-1 and HIV-2 types **(Middle)** [390]. Dashes indicate sequence identity to HIV-1 consensus sequence. Importance of each individual amino acid of N36 for HIV-1 infection, according to Sen et al[372] **(Bottom)**. In the cited study, an alanine substitution was performed for each amino acid of N36 region, classifying each one according to the percentage of HIV-1 infection comparing to the wild-type. Wild-type mutants: >40% viral entry; Impaired entry-mutants: 5-40% viral entry and non-function entry mutants: <5% viral entry.

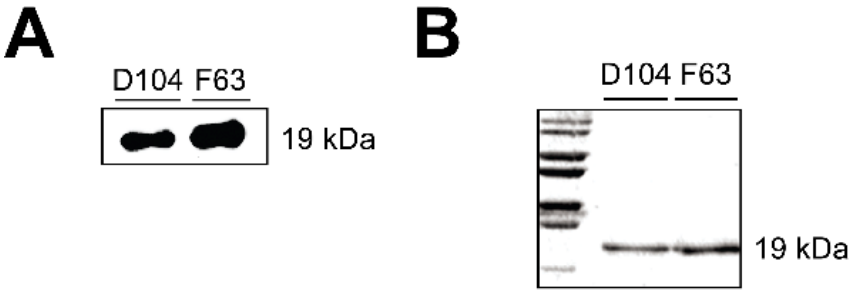


Figure 2.9: Analysis of purified VLs D104 and F63.
Western-blot **(A)** and SDS-PAGE **(B)** of purified VLs D104 and F63.

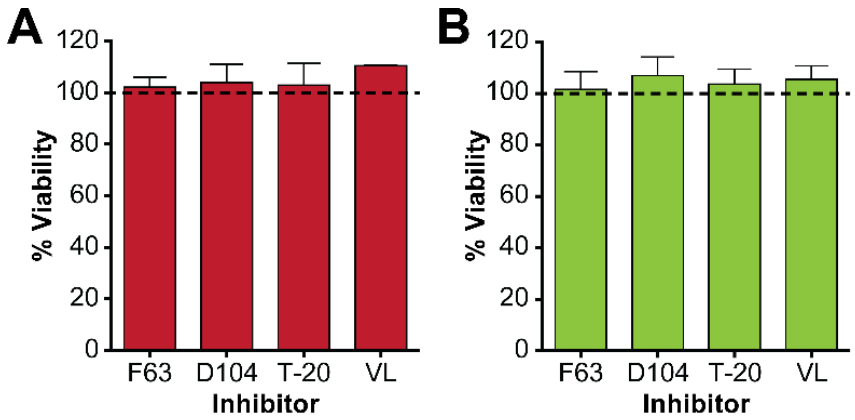


Figure 2.10: Viability assays.
TZM-bl **(A)** and Jurkat cells **(B)** viability was determined in the presence of the highest used concentration of the inhibitors (F63, D104, T-20 and VL). Error bars correspond to standard deviation (n = 3).

2.8.2 Supplementary Tables

Table 2.1: PCR fragments and primers used in this study.

VL CDR3 22 aa F

> VL-F-SfiI

5'- GAGGAGGAGGAGGAGGAGGCGGGGCCCAGGCGGCCGAGCTC -3'

> VL Framework 3 22 aa R

5'-CCTCAATTCG TAGTAGCTA GCGTAGATG GTGGAGTCG GAGCCAC CGCCTGA ACAGTAGT AAGTGGCAGCATCGGC -3'

VL CDR3 22 aa R

> VL CDR3 22 aa FF

5'-GACT CCACCAT CTACGCT AGCTACTA CGAATTG AGGAGCG GAGGCG GAAGTT TTGC TTTCGGCGGAGGG-3'

> VL-R-SfiI

5'-CCGCTCGA GCGGCTAAG AAGCGTAGTC CGGAACGTCG TACGGGTAAG AAGCGTAGTC CGGAACGTC-3'

VL CDR3 26 aa F

> VL-F-SfiI

5’-GAGGAGGAGGAGGAGGAGGCGGGGCCCAGGCGGCCGAGCTC-3’

> VL Framework 3 26 aa R

5’-CCTCAATTCGT AGTAGCTAGCG TAGATGGTGGA GTCGGAGCC ACCGCCAC CGCCTGAC CCGGAAC AGTAGT AAGTGGCAGCATCGGC-3’

VL CDR3 26 aa R

> VL CDR3 26 aa FF

5’-GACTC CACCATC TACGCTA GCTACT ACGAATT GAGGAG CGGAGG CGGAAG TGGT TCATTTG CTTTCGGCGGAGGG-3’

> VL-R-SfiI

5’-CCGCTC GAGCGGCTAAG AAGCGTAGT CCGGAACGTCG TACGGGTAAG AAGCGTAGTCC GGAACGTC-3’

VL CDR3 30 aa F

> VL-F-SfiI

5’-GAGGAGGAGGAGGAGGAGGCGGGGCCCAGGCGGCCGAGCTC-3’

> VL Framework 3 30 aa R

5’-CCTCAATTC GTAGTAGCTA GCGTAGAT GGTGGAGTCG GAGCCACCGCC AGAGCCTCCC CCGGAACAG TAGTAAGTGGCAGCATCGGC-3’

VL CDR3 30 aa R

> VL CDR3 30 aa FF

5’-GACTC CACCATC TACGCTA GCTACT ACGAATT GAGGAG CGGAGG CGGAAG TGGT TCATTTG CTTTCGGCGGAGGG-3’

> VL-R-SfiI

5’-CCGCTCGA GCGGCTAAG AAGCGTAGTC CGGAACGTC GTACGGGT AAGAAGCG TAGTCCGG AACGTC-3’

HEL binders overlap

> VL-F-SfiI

5’-GAGGAGGAGGAGGAGGAGGCGGGGCCCAGGCGGCCGAGCTC-3’

> VL-R-SfiI

5’-CCGCTCGAGCG GCTAAGA AGCGTAGTCC GGAACGTCGT ACGGGTAAG AAGCGTAGTCC GGAACGTC-3’

CDR1 randomization F

> VL-F-SfiI

5’-GAGGAGGAGGAGGAGGAGGCGGGGCCCAGGCGGCCGAGCTC-3’

> CDR1-R

5’-ACTTCCGCCTCCGCTGCAATTGATGGTGACTGTGCCC-3’

CDR1 randomization R

> CDR1 library F

5’-CAATTGCAGCGGAGG CGGAAGTD VNDVNDVNDVNDVNDVN DVNDVNDVNDVNDVN DVNAGCGGAGGCGGAAG TTGGTATCAGCAGAAACCAGGGC-3’

> VL-R-SfiI

5’-CCGCTCGAG CGGCTAAG AAGCGTAGT CCGGAACGTC GTACGGGT AAGAAGCG TAGTCC GGAACGTC-3’

CDR1 randomization overlap

> VL-F-SfiI

5’-GAGGAGGAGGAGGAGGAGGCGGGGCCCAGGCGGCCGAGCTC-3’

> VL-R-SfiI

5’-CCGCTCG AGCGGCTA AGAAGCGTAG TCCGGAACGTC GTACGGGTAA GAAGCGTAGTCC GGAACGTC-3’

CDR3 randomization F

> **VL-F-SfiI**

5’-GAGGAGGAGGAGGAGGAGGCGGGGCCCAGGCGGCCGAGCTC-3’

> **CDR3-R**

5’-ACTTCCGCCTCCGCTACAGTAGTAAGTGGCAGCATCGGC-3’

CDR3 randomization R

> **CDR3 library F**

5’-CTTA CTACTGTAG CGGAGGCGGAAG TDVNDVNDVND VNDVNDVNDVN DVNDVNDVND VNDVNAGCGGAGGCGGAAGTTTTGCTTTCGGCGGAGGG-3’

> **VL-R-SfiI**

5’-CCGCT CGAGCGGCTA AGAAGCGTAGTC CGGAACGTCGTAC GGGTAAGAAGCGTA GTCCGGAACGTC-3’

VL NheI/XhoI for pET28a (+)

>VL-NheI-F

5’- CTC GCT AGC GAG CTC GTG CTG ACC CAG-3’

> VL-XhoI-R

5’-CCG CTC GAG GCT GCC TCC GCC TCC GCT TAG GAT CTC CAG CTC GGT CCC-3’

F63 monomer N-terminal

>VL-NheI-F

5'- CTC GCT AGC GAG CTC GTG CTG ACC CAG-3'

>VL-NotI-R linker

5'-TTTTCCTTTTGCGG CCGCTGCTACCTCCA CCTCCGCTGTGCTG GGCGGCCTGGCCA GGCCAGGCCGCCCA GCACAGCGG AGGTGGAGGTAGCAGCGGCCGCAAAAGGAAAA-3'

F63 monomer C-terminal

>VL-NotI-F

5'-AAG GAA AAA AGC GGC CGC GCC CAG GCG GCC GAG CTC-3'

> VL-XhoI-R

5'-CCG CTC GAG GCT GCC TCC GCC TCC GCT TAG GAT CTC CAG CTC GGT CCC-3'

Table 2.2: Binding analysis of phage-selected VLs against N36.

	EC_{50}[a] ±SD[b], nM
C62	266.5 ±1.07
D103	284.6 ±1.08
F63	274.5 ±1.08
D104	234.6 ±1.12
G54	287.62 ±1.11
VL[c,d]	n.a.

[a]EC_{50} (50% effective concentration) were inferred from sigmoidal dose-response (variable slope) curves.

[b]SD stands for standard deviation of n = 3.

[c]VL represents the $VL_{parental}$.

[d]n.a. stands for not applicable.

Table 2.3: Antiviral activity of F63 against HIV-1 primary isolates in PBMCs.

HIV-1 primary isolate[a]	Subtype	F63 IC_{50}[b] ±SD[c], nM	T-20 IC_{50}[b] ±SD[c], nM	VL[d,e]
UCFL1014	B	4.2 ±0.4	1.2 ±0.3	n.a.
UCFL1025	B	0.4 ±0.1	0.2 ±0.1	n.a.
UCFL1028	C	0.3 ±0.1	0.1 ±0.1	n.a.
UCFL1029	B	7.0 ±2.0	2.5 ±0.8	n.a.

[a]The HIV-1 primary isolates were obtained from Calado et al [355].

[b]IC_{50} (50% inhibitory concentration) were inferred from sigmoidal dose-response (variable slope) curves.

[c]SD stands for standard deviation of n = 3.

[d]VL represents the $VL_{parental}$.

[e]n.a. stands for not applicable.

Table 2.4: Membrane partition constants of F63.

	F63 $K_p \times 10^3$ ±SD[a]	VL[b] $K_p \times 10^3$ ±SD[a,c]
POPC	3.27 ±0.70	~0
POPC:Chol (2:1)	3.49 ±1.06	~0

[a]SD standard deviation of $n = 3$.

[b]VL represents the $VL_{parental}$.

[c]Data fitting was not possible.

Table 2.5: Binding kinetics of F63.

k_a ±SD[a], M^{-1} s^{-1}	k_d ±SD[a], s^{-1}	$K_D = k_d/k_a$ ±SD[a], M	K_D ±SD[a], nM
6.1×10^4 ±1.2	5.3×10^{-4} ±0.3	8.8×10^{-9} ±1.1	8.8 ±1.1

[a]SD standard deviation of n = 3

CHAPTER III

CXCR4-TARGETED NANOBODY FOR DELIVERY OF MODULATORS OF HIV GENE EXPRESSION

Catarina Cunha-Santos,[1] and Joao Goncalves[1]

[1] *Research Institute for Medicines (iMed.ULisboa), Faculty of Pharmacy, Universidade de Lisboa, Lisbon, Portugal*

MANUSCRIPT IN PREPARATION

CXCR4-TARGETED NANOBODY FOR DELIVERY OF MODULATORS OF HIV ENE EXPRESSION

3.1 Abstract

Small interfering RNAs (siRNA) application in therapy still faces a major challenge with the lack of an efficient and specific delivery system. Here, we present a novel strategy for sucessfully deliver of an anti-Human Immunodeficiency Virus (HIV) siRNA. Membrane translocation of RNAi effector was addressed by an engineered nanobody towards CXCR4 receptor, a major target expressed on HIV-susceptible T-lymphocytes. A validated siRNA molecule against HIV-1 gene *tat* was conjugated to fluorescein (FITC) and coupled to the CXCR4-specific nanobody through an anti-FITC single-chain variable fragment (4M5.3 scFv). We chose a high-affinity binder of FITC—4M5.3 affinity constant in the fentomolar range—to minimize siRNA loss. The siRNA-fusion protein targeted CXCR4 positive cells, being detected at the surface and in the cytoplasm of a human T-lymphocytic cell line. Additionally, our construct silenced expression of *luciferase* reporter gene under control of Tat-driven HIV promoter and inhibited HIV-1 infectivity. Similarly to RNA interference technology, zinc-finger transcription factors are potent modulators of gene expression. Accordingly, we further demonstrate the suitability of this nanobody as a vehicle for CXCR4 entry-dependent functionality of a delivered anti-HIV zinc-finger repressor. The present study demonstrates the potential of a specific delivery system for therapeutic HIV molecules based on nanobody chimeras towards CXCR4 receptor.

3.2 Introduction

Methods of gene modulation have gained tremendous interest for therapeutic purposes. One of most well-known mechanisms is the RNA interference (RNAi). This technology of post-transcriptional gene silencing is mediated by RNA duplexes of 21-3 nucleotides, termed small interfering RNAs (siRNAs), that trigger the cellular degradation of cognate mRNAs.[224] RNAi silencing potency, specificity of target, universal conservation and broad application, namely to non-druggable targets, render it attractive for therapeutics. However, non-efficient cellular uptake of siRNAs, derived from their intrinsic negative charge and hydrophilicity, settles the need for delivery strategies that enable these RNA-based molecules translocation across cellular membrane.

Viral vectors and non-virus systems such as liposomal and polymeric nanocarriers, cell-penetrating peptides, ligands of cell receptors or antibody-based formats directed to cell-surface proteins are amongst the most widespread carriers of RNAi effectors.[245,257,391] Cell-targeted delivery presents several advantages such as reduction of required drug amount and prevention of off-target silencing and toxicity effects besides an overall potentiation of the therapeutic benefit. Nonetheless, and despite the diverse available methods, some limitations have been associated with systemic delivery of therapeutic siRNAs. For example, conjugation of these RNAi effectors with small size ligands such as aptamers or cell-penetrating peptides does not prolong short-term circulation due to threshold of renal excretion.[279] Liposome-based methods are also known to easily trigger diverse inflammatory pathways[261] and to hamper intracellular release of siRNA molecules.[256] In other cases, a high dosage is required to induce a therapeutic benefit,[236] which results in burden costs, possible off-target effects and toxicity.

Recent clinical trials reported tumor regression in patients treated with siRNA-based drugs delivered by lipid carriers.[392,393] Regardless of these advances, siRNA therapeutics in non-oncogenic diseases such as HIV infection remains a hurdle mainly due to lack of tumoral enhanced permeability and retention effect.[394,395] Moreover, HIV-target lymphocytes are not only described as hard-to-transfect cells with conventional lipid-based strategies but also as being concentrated on densely packed environments, limiting the size of the chosen carrier. HIV co-receptor CXCR4 is overexpressed in $CD4^+$ T lymphocytes, main target cells of HIV, being efficiently internalized shortly after interaction with its natural ligand, SDF-1α.[396] Previous studies have successfully

implemented a CXCR4-binding antibody for targeted delivery of cytotoxic drugs for cancer cells treatment[397] as well as CXCR4-specific peptides for transfection of DNA molecules.[398,399] Accordingly, we explored the potential of CXCR4 targeting to develop a novel strategy for antibody-based delivery of anti-HIV siRNAs.

Here, we report the delivery of a siRNA targeting viral *tat* transcript via a scFv (single-chain variable fragment)-nanobody fusion protein, herein termed 4M5.3C. We show proof-of-concept of 4M5.3C capacity to provide selective binding and consequently internalization of siRNA cargo into CXCR4-bearing cells. Our results prove that 4M5.3C-mediated delivery of siRNA silences HIV transactivator Tat expression and inhibits viral infectivity in cultured T-cells. We further demonstrate the use of this anti-CXCR4 nanobody for targeted delivery of engineered zinc-finger repressors of HIV expression. In this approach, we show that zinc-finger inhibition activity is dependent on CXCR4-mediated endocytosis as delivery route. Both strategies support the concept of using nanobody-based constructions for targeted delivery of repressors of HIV expression. To the best of our knowledge, we demonstrate for the first time the potential of nanobody-based vehicles for efficient delivery of HIV inhibitor molecules.

3.3 Materials and Methods

3.3.1 Cells and viruses

The following reagents were obtained through the NIH AIDS Reagent Program (Division of AIDS, NIAID, NIH): HeLa-*tat*-III from Drs. William Haseltine, Ernest Terwilliger and Joseph Sodroski,[400,401] HeLa-Tat-III/LTR/d1EGFP from Dr. Satoh,[402] Jurkat Clone E6-1 T cells from A. Weiss,[389] Sup-T1 from James Hoxie,[403] TZM-bl from J. C. Kappes, X. Wu and Tranzyme Inc.,[163,384–387] and pNL4-3 from M. Martin.[366] HEK293T cells were purchased from ATCC.

All cell cultures were maintained at 37 °C in a humidified atmosphere of 5% CO_2. Jurkat and Sup-T1 cells were maintained in RPMI-1640 medium (Lonza) supplemented with 10% (v/v) fetal bovine serum (FBS; Thermo scientific), 1% (v/v) antibiotic/antimycotic solution (PSA; Thermo scientific) and 2 mM L-glutamine (Thermo scientific). HEK293T, HeLa-*tat*-III, HeLa-Tat-III/LTR/d1EGFP and TZM-bl were maintained in DMEM medium (Lonza) supplemented with 10% (v/v) FBS, 1% (v/v) PSA and 2 mM L-glutamine. In addition, HeLa-*tat*-III and HeLa-Tat-III/LTR/d1EGFP were kept in 1 mg/mL geneticin (Gibco).

3.3.2 siRNAs

siRNA directed against *tat* was as previously described[404]:

5'-GCGGAGACAGCGACGAAGAGCTTdTdT-3' (sense);

5'-GCUCUUCGUCGCUGUCUCCGCdTdT-3' (anti-sense).

The siRNA described was labeled with FITC at the 5'end of the sense strand (A4 grade, Dharmacon Research). Scrambled siRNA (MISSION siRNA Universal Negative Control #1 6-FAM) was purchased from Sigma-Aldrich (Germany).

3.3.3 Constructions

FITC-binding scFv (4M5.3),[405] CXCR4-targeting (282D2)[406] and irrelevant nanobodies were synthetized by Geneart (Germany). Sequence of CXCR4-targeting 282D2 nanobody was obtained from US 2011/0318347 A1 patent[407] and sequence of KRAB-HLTR3[333] from GenBank (Accession number AY518587.1). Fragments of 4M5.3 scFv, KRAB-HLTR3 and CXCR4-specific and irrelevant nanobodies were originated by PCR using the primer sequences provided in Supporting Information (Table 3.1). 4M5.3C, 4M5.3I, KRAB-HLTR3C and KRAB-HLTR3I constructions were assembled by ligation and cloned into the NheI and XhoI restriction sites of pET-21a(+) expression vector (Novagen). CXCR4-targeting nanobody and KRAB-HLTR3 alone were directly cloned into NheI/XhoI sites of pET-21a(+) plasmid. Constructions were verified by DNA sequencing analysis.

3.3.4 Protein expression and purification

Nanobodies for targeted siRNA delivery were produced by soluble protein purification in *Escherichia coli* shuffle cells (NEB). Chemically competent cells were transformed with 4M5.3C, 4M5.3I and anti-CXCR4 nanobody alone. Cultures were growth in 1 L of Super Broth (SB) medium supplemented with 100 µg/mL of ampicillin and 20 mM of $MgCl_2$ at 30 °C until OD_{600} of 0.4, induced with 0.4 mM IPTG and further incubated at the same temperature for 4 h. Bacterial cells were collected and resuspended in 40 mL of 20 mM HEPES pH 7.4, 500 mM NaCl, and 20 mM imidazole before disruption through sonication during 20 min. Proteins were purified by IMAC, using His GraviTrap columns (GE Healthcare). Purified proteins were buffer exchanging to 20 mM HEPES pH 7.4, 500 mM NaCl and 5% glycerol using PD-10 Desalting Columns (GE Healthcare).

Nanobodies for targeted zinc-finger delivery were produced by insoluble protein purification as previously described.[408] Chemically competent *E. coli* BL21 (DE3) cells

(Stratagene) were transformed with constructions KRAB-HLTR3, KRAB-HLTR3C and KRAB-HLTR3I. Cultures were growth in lysogeny broth (LB) medium supplemented with 100 µg/mL of ampicillin at 37 °C until OD_{600} of 0.9, induced with 1 mM IPTG and further incubated at the same temperature for 4 h. Bacterial cells were collected and resuspended in 40 mL of 50 mM HEPES pH 8.0, 1 M NaCl, 10 mM imidazole, 2 M urea and 5 mM $CaCl_2$ before disruption through sonication. The insoluble fraction was cleared by centrifugation, resuspended in the same buffer and further disrupted by sonication. Finally, insoluble proteins were resupended in 50 mM HEPES pH 8.0, 1 M NaCl, 10 mM imidazole, 6 M urea, 1 mM 2-mercaptoethanol, 5 mM $CaCl_2$ and solubilized overnight at 4 °C with gentle agitation. Soluble denatured proteins were recovered after centrifugation and purified by IMAC, using His GraviTrap columns (GE Healthcare). Purified proteins were renatured by buffer exchanging to 20 mM Tris-HCl pH 7.4, 1 mM $MgCl_2$, 90 mM KCl and 0.1 mM $ZnCl_2$ using PD-10 Desalting Columns (GE Healthcare).

Purification of all proteins was confirmed by Western-blot technique as described previously.[408] Purity was assessed by Coomassie-blue staining and quantification measured by Bradford method according to manufacturer instructions.

3.3.5 Cell lines construction and virus production

The protocol for construction of Jurkat $CXCR4^-$ cell line through CRISPR/Cas9 gene knockout was adapted from Hou et al.[409]. For production of CXCR4-CRIPR/Cas9 lentivirus, HEK293T seeded in 6-well plate were transfected with 1.3 µg lentiCXCR4-gRNA-Cas9 #6, 1 µg psPAX2 and 0.7 µg pMD2.G plasmids (kindly provided by D. Guo) using lipofectamine 3000 reagent (Invitrogen) according to the improved lentiviral production protocol from the manufacturer. Jurkat T-cells ($1.0x10^5$) seeded in 24-well plate were transduced with 400 ng of p24 and incubated for 3 days. Receptor negative cells were sorted by flow cytometry, after negative selection of CXCR4-stained cells with APC (allophycocyanin) conjugated anti-human CXCR4 antibody 12G5 (Biolegend).

To produce HIV viral particles, HEK293T seeded in 6-well plate were transfected with 3 µg pNL4-3 plasmid using lipofectamine 3000 reagent (Invitrogen) according to the improved lentiviral production protocol from the manufacturer.

All viral titers were determined by HIV-1 $p24^{CA}$ antigen capture assay kit (Frederick National Laboratory for Cancer Research – AIDS and Cancer Virus Program).

3.3.6 Cell-surface binding and internalization of therapeutic siRNA

FITC-conjugated scramble and *tat* siRNAs were mixed with 4M5.3C, 4M5.3I control or PBS at a molar ratio 2:1 (siRNA 200 pmol) in PBS for 1 h at 37 °C. $CXCR4^+$ or $CXCR4^-$ Jurkat T-cells ($2.0x10^5$ in 100 µl culture medium) were treated for 2 h at 4 °C or 37 °C, washed twice with PBS or trypsin for 3 min, respectively, and analyzed by flow cytometry. For positive control, Jurkat cells were transfected with FITC-siRNA using interferin (Polyplus) according to manufacturers' instructions. For competition assays, *tat* siRNA was mixed with 4M5.3C in the presence of crescent quantities of VHH α-CXCR4 (competitor) for 2 h at 37 °C. Cells were washed twice with trypsin for 3 min and analyzed by flow cytometry for detection of FITC-positive population.

3.3.7 Western-blot zinc-finger transcription factors

Zinc-fingers detection in cells cytoplasm was adapted from Perdigão et al.[410] Briefly, ZF-TF proteins were diluted in RPMI medium containing 100 µM $ZnCl_2$ and added to Jurkat cells ($5.0x10^5$) seeded in 6-well plate. Cells were treated for 2 h, 20 h and 24 h. After treatment, cells were harvested and washed twice with trypsin for 3 min on ice to eliminate protein surface-binding. Cells were lysed with RIPA buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate and 0.1% SDS) supplemented with EDTA-free Protease Inhibitor Cocktail Tablets (Roche, Basel, Switzerland) and TFs entry analyzed by SDS-PAGE. Samples were transferred onto a 0.2 µm nitrocellulose membrane and detected with Immobilon Western Chemiluminescent HRP substrate (Millipore, Billerica, MA, USA) and Amersham HyperfilmECL (GE Healthcare, Little Chalfont, UK) chemiluminescence film. Zinc-finger repressors were detected by horseradish peroxidase-conjugated anti-HA monoclonal antibody (Roche, Basel, Switzerland). β-actin was used as an internal control and detected by a mouse anti-β-actin monoclonal antibody (Sigma, St. Louis, MO, USA) and horseradish peroxidase-conjugated goat anti-mouse IgG antibody (Bio-Rad, Hercules, CA, USA).

3.3.8 Silencing/Shutdown of HIV LTR expression

For siRNA-mediated LTR shutdown, TZM-bl cells were plated onto 24-well plates at a density of $1.0x10^5$ cells per well. At 24 h after plating, cells were transfected with 500 ng of pTat plasmid (kindly provided by M. Simurda) using lipofectamine 3000 (Invitrogen) according to manufacturer's instructions. At 24 h post-transfection, scramble and *tat* siRNAs were mixed with 4M5.3C, 4M5.3I control or PBS at a molar ratio 2:1 (siRNA 200 pmol) in PBS for 1 h at 37 °C. TZM-bl cells were then washed with PBS

and treated with these siRNA carriers for 3 h. At 24 h post-treatment, luciferase activity was measured by OneGlo^{TM} Luciferase Assay System (Promega) according to manufacturer´s instructions. Luciferase activity was normalized to total amount of protein, as quantified by Bradford method. Cells' viability was assessed by AlamarBlue reagent according to manufacturer's instructions.

For zinc-finger-mediated LTR shutdown, HeLa-Tat-III/LTR/d1EGFP cells were seeded into 24-well plates at 1.0x10^5 of confluence. After 24 h, cells were treated with 200 nM of zinc-finger repressors and nanobody control for 3 h. Afterwards, cells were washed twice with PBS and analyzed by flow cytometry to evaluate GFP expression and exclude apoptotic cells after staining with Alexa 405-annexin V.

3.3.9 HIV inhibition assays

For siRNA-mediated viral inhibition, Sup-T1 cells (1.0x10^5) were infected with 10 ng p24 of HIV-1 NL4-3 in 24-well plates. At 24 h post-infection, scramble and *tat* siRNAs were mixed with 4M5.3C, 4M5.3I control or PBS at a molar ratio 2:1 (siRNA 200 pmol) in PBS for 1 h at 37 °C and further incubated for with Sup-T1 cells 3 h. After 3 days of infection, HIV replication was assessed by p24 capsid quantification on culture supernatants (HIV-1 p24^{CA} antigen capture assay kit; Frederick National Laboratory for Cancer Research – AIDS and Cancer Virus Program, Frederick, MD).

3.4 Results

3.4.1 CXCR4-targeted siRNA delivery by 4M5.3C

We took advantage of a previously validated CXCR4-targeted heavy-chain variable domain from camelid (nanobody)[406] and explored its potential to promote endocytosis-mediated delivery of a siRNA inhibitor. CXCR4-targeted nanobody, herein named NbCXCR4, was fused to the C-terminus of an anti-FITC scFv though the flexible linker GGGGS to originate the 4M5.3C chimera (Fig. 3.1A). The 4M5.3 scFv was chosen due to its exceptionally binding affinity for fluorescein (FITC)[405] in the femtomolar range to carry the fluorochrome-conjugated siRNA. We also constructed a fusion protein composed of 4M5.3 portion and an irrelevant nanobody (non-targeting for mammalian cells) and the NbCXCR4 alone as negative controls for targeted delivery. All constructions were expressed and purified from *E. coli*, being purity assessed as >90% by SDS-PAGE (Fig. 3.1B). Proteins purification was also confirmed by Western-blot (data not shown).

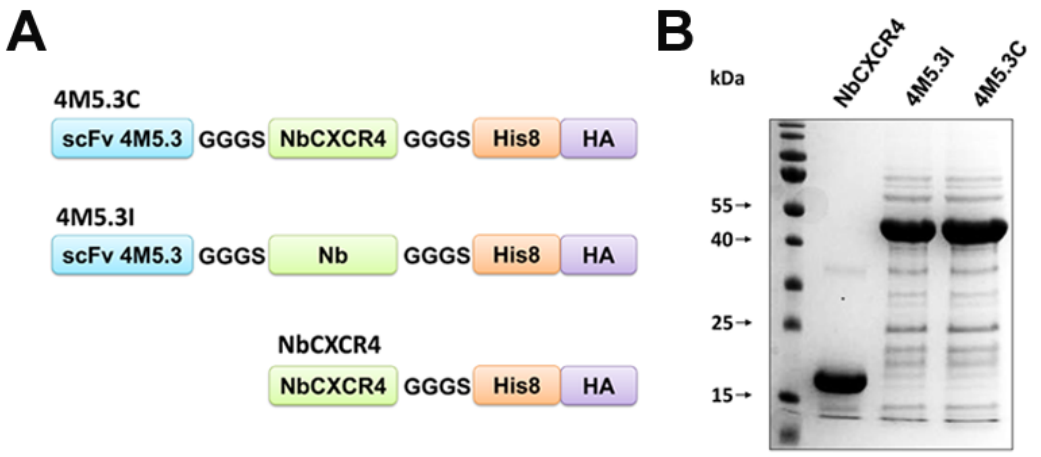


Figure 3.1: Design of scFv-nanobody chimera for targeted delivery of FITC-conjugated siRNAs.
A) Schematic representation of siRNA delivery constructs. Anti-FITC scFv (4M5.3) is positioned at the N-terminal of CXCR4-targeted (NbCXCR4) or irrelevant (Nb) nanobody to generate 4M5.3C or 4M5.3I, respectively. NbCXCR4 control construct is devoid of anti-FITC scFv fragment. GGGS linkers were placed between the scFv fragment and the nanobody portion or histidine (His) and hemagglutinin A (HA) tags, respectively for protein purification or detection. **B)** SDS-PAGE analysis of purified protein constructs. Proteins were detected by Coomassie-blue staining.

To evaluate receptor-specific internalization of siRNA, $CXCR4^+$ or $CXCR4^-$ Jurkat T-cells were incubated with FITC-siRNA alone or in complex with 4M5.3C (2:1 molar ratio). CXCR4-knockout cell line was constructed using the lentiCRISPR/Cas9 technology for disruption of CXCR4 gene adapted from Hou *et al*[409] (Fig. 3.2A). FITC-siRNA lipid-based transfection of Jurkat T-cells was used as positive control for delivery. FITC fluorescence was only detected in $CXCR4^+$ Jurkat T-cells incubated with 4M5.3C+siRNA conjugate (Fig. 3.2B). As expected, FITC-labeled siRNA alone did not appreciably enter both cell lines. 4M5.3C-mediated internalization of CXCR4 appears to have no deleterious effect on cells viability. We observed that increase of FITC fluorescence in $CXCR4^+$ Jurkat T-cells is dose-dependent of 4M5.3C+siRNA concentration (Fig. 3.2C). To further exclude a possible unspecific cell-permeability of 4M5.3C, we compared FITC-siRNA binding (4ºC) and internalization (37ºC) between CXCR4-targeted 4M5.3C and irrelevant 4M5.3I control. As showed in Fig 3.2D, only 4M5.3C construct internalized the FITC-tagged siRNA into the $CXCR4^+$ Jurkat T-cells (% FITC was 44.21 ±2.00 for 4M5.3C+siRNA vs 5.14 ±0.87 for 4M5.3I+siRNA; $p<0.001$). A similar effect was observed with CXCR4 binding at cell-surface, further verifying 4M5.3C specificity to CXCR4. To further evidence 4M5.3C specificity towards target receptor, we incubated $CXCR4^+$ Jurkat T-cells with FITC-siRNA/4M5.3C complexes in the presence of NbCXCR4 alone as competitor for CXCR4 binding (Fig. 3.2E). We observed that detection of FITC decreased as the concentration of NbCXCR4 competitor increased, demonstrating that blocking CXCR4 receptor abolishes siRNA

internalization by 4M5.3C. Overall, these results indicate that engineered 4M5.3C can efficiently deliver FITC-conjugated siRNA through CXCR4-endocytosis.

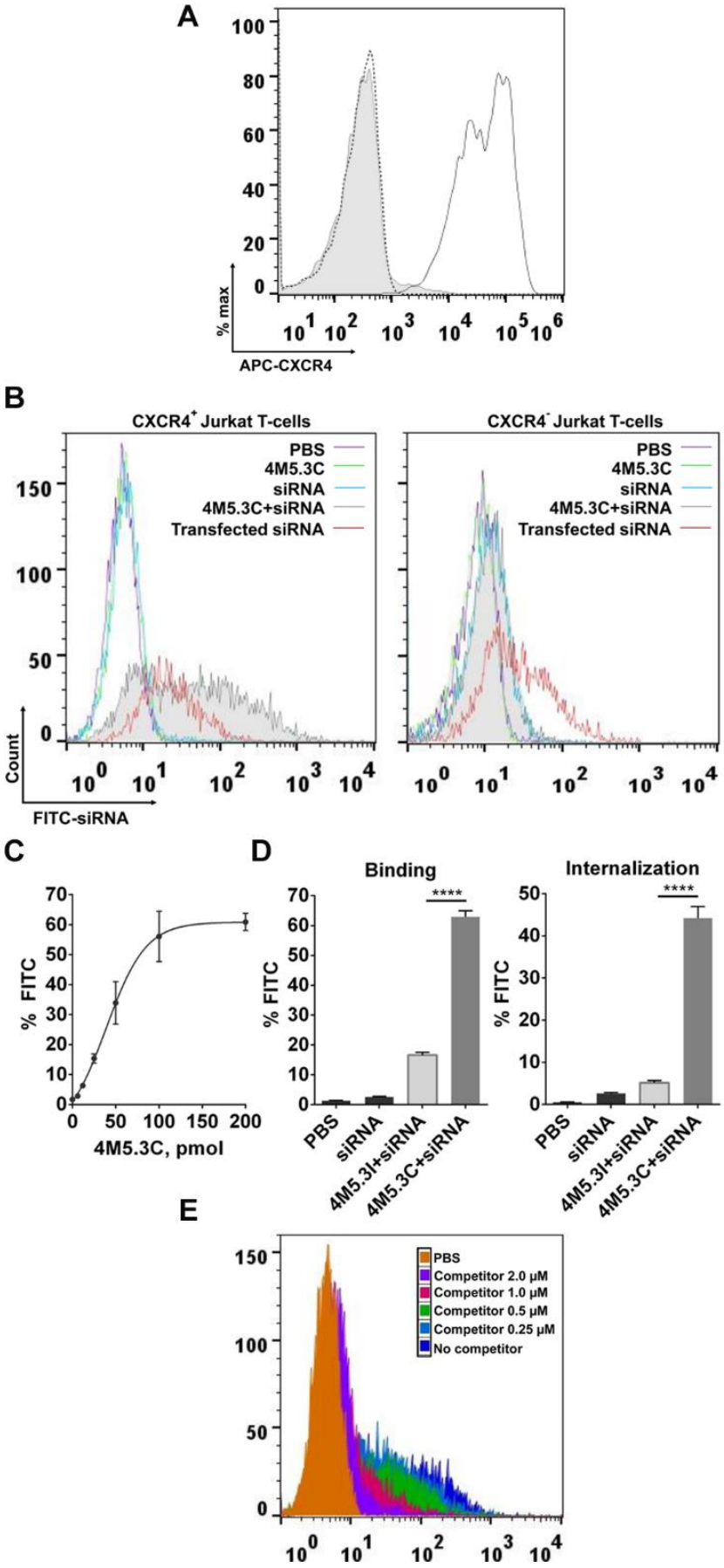


Figure 3.2: CXCR4-targeted delivery of FITC-conjugated siRNA through scFv-nanobody chimera.

A) Histogram illustrating CXCR4 expression of CXCR4-positive or CXCR4-negative Jurkat T-cells stained with APC-conjugated anti-CXCR4 and analyzed by flow cytometry. The solid black line represents the $CXCR4^+$ Jurkat T-cells stained with anti-CXCR4 antibody. The dotted black line represents the $CXCR4^+$ Jurkat T-cells unstained. The grey color represents the $CXCR4^-$ Jurkat T-cells stained with CXCR4-specific antibody. **B)** Histograms illustrating internalization of FITC-conjugated siRNA in CXCR4-positive or CXCR4-negative cells treated with 4M5.3C or siRNA alone, or 4M5.3C-siRNA complex (2:1 ratio) for 2 h at 37ºC. Cells were washed with trypsin to eliminate the CXCR4-surface bound molecules before FITC detection by flow cytometry. "PBS" represents untreated cells. "Transfected siRNA" represents lipid-based siRNA transfection, the positive control for delivery. **C)** Percentage of $FITC^+$ cells following flow cytometry analysis of CXCR4-positive Jurkat treated with increasing amounts of 4M5.3C-siRNA conjugation for 2 hours at 37ºC. **D)** Percentage of $FITC^+$ cells following flow cytometry analysis of CXCR4-positive Jurkat treated with CXCR4-targeted 4M5.3C-siRNA of irrelevant 4M5.3I-siRNA conjugations for 2 h at 4 ºC (binding) or 37 ºC (internalization). Only cells incubated at 37 ºC were washed with trypsin. "PBS" represents untreated cells. "siRNA" indicates cells incubated with FITC-siRNA alone. **E)** Histogram illustrating FITC-siRNA internalization in CXCR4-positive Jurkat cells treated with 4M5.3C-siRNA in the absence (No competitor) or presence of increasing concentrations of NbCXCR4 competitor. "PBS" represents the untreated cells. For panels **a**, **b** and **e**, histograms represent one of at least 3 independent assays. For panel **c**, values represent mean ± SEM of at least 6 independent assays. For panel **d**, values represent mean ± SEM of at least 10 independent assays. 4M5.3C+siRNA vs 4M5.3I+siRNA ****p<0.0001 (Mann-Whitney test).

3.4.2 4M5.3C-mediated siRNA delivery silences Tat-driven HIV transcription

To evaluate whether 4M5.3C-delivered siRNA could silence target gene expression, $CXCR4^+$ TZM-bl cells were transfected with pTat plasmid and afterwards incubated with 4M5.3C fusion protein carrying the *tat* siRNA. These cells encode the *luciferase* reporter gene under control of the long-terminal repeat (LTR)—HIV promoter—which in turn is transactivated by the Tat viral protein. Lipid-based transfection of *tat* siRNA was used as positive control for silencing of LTR expression. As demonstrated in Fig. 3.3A, siRNA delivered by 4M5.3C reduced LTR transcription by ~85% ($p<0.05$). In contrast, the presence of 4M5.3C+siRNA scramble (control) did not significantly affect *luciferase* transcription from the LTR promoter. We confirmed that 4M5.3C-mediated internalization of this non-targeting siRNA was similar to the anti-*tat* siRNA (data not shown). Furthermore, silencing of LTR expression in the presence of 4M5.3C+siRNA was not due to a loss in cell viability as demonstrated in Fig. 3.3B. These data indicate that delivery of anti-Tat siRNA through 4M5.3C construct specifically silences Tat-driven LTR transcription.

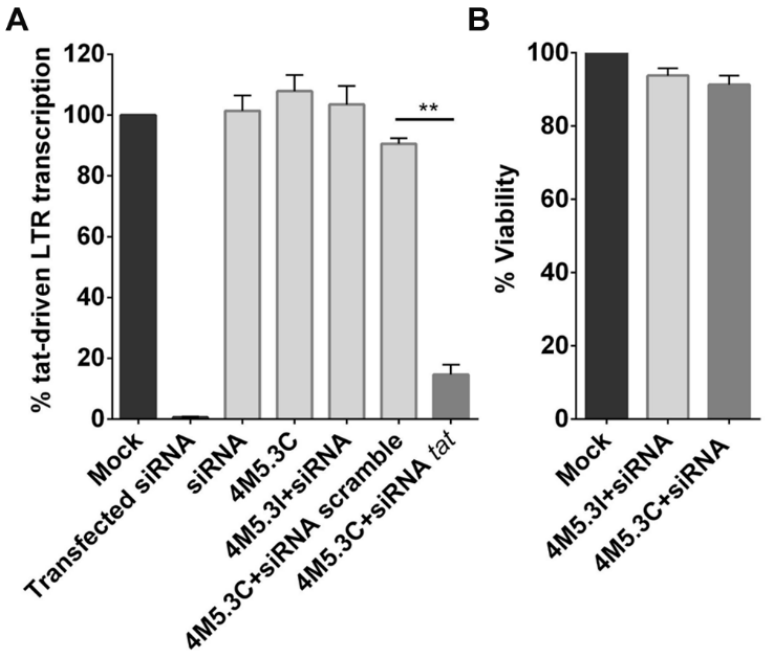


Figure 3.3: tat siRNA delivered by 4M5.3C silences Tat-driven LTR transcription in HIV reporter cells.
A) Evaluation of Tat-driven LTR transcription assessed by luciferase activity in TZM-bl reporter cell line transfected with pTat plasmid followed by treatment with FITC-conjugated tat siRNA (siRNA), 4M5.3C alone or 4M5.3C-siRNA conjugation. Cells were transfected with pTat 24 hours previously to protein treatment. Cells were also incubated with FITC-tagged siRNA scramble conjugated with CXCR4-targeted 4M5.3C construct (4M5.3C+siRNA scramble) or irrelevant 4M5.3I control conjugated with tat siRNA (4M5.3I+siRNA). "Mock" represents untreated cells. "Transfected siRNA" represents lipid-based siRNA transfection for positive control of LTR transcription silencing. **B)** Viability of TZM-bl cells in presence of tat siRNA-conjugated 4M5.3C and 4M5.3I constructs. Cell viability was assessed by Alamar Blue viability assay 3 h after treatment. Values represent mean ± SEM of at least 3 independent assays and are normalized to pTat transfected cells. 4M5.3C+siRNA vs 4M5.3I+siRNA or vs 4M5.3C+siRNA scramble**$p<0.01$ (Mann-Whitney test).

3.4.3 4M5.3C-mediated siRNA delivery inhibits HIV replication

For the therapeutic siRNA to exert its antiviral activity, it is essential that 4M5.3C allows CXCR4-mediated internalization into HIV-infected cells. To evaluate inhibitory effect of CXCR4-targeted delivery of anti-Tat siRNA, a T-lymphotropic cell line was infected with HIV-1 laboratory-adapted strain NL4-3 (CXCR4-tropic) and afterwards incubated with 4M5.3C+siRNA conjugates. 4M5.3C loaded with *tat* siRNA reduced HIV replication on ~80% in T-lymphotropic cell line (Fig. 3.4A). We observed partial HIV inhibition with 4M5.3C alone, possibly due to NbCXCR4 reported viral neutralization activity as a fusion inhibitor.[406] Nevertheless, HIV infection was dramatically decreased only when this construct was conjugated with anti-Tat siRNA. As demonstrated in Fig. 3.4B, neither 4M5.3C conjugated with scramble siRNA nor *tat* siRNA conjugated with irrelevant 4M5.3I were able to inhibit NL4-3 infection. In conclusion, our results demonstrate that CXCR4-targeted delivery of anti-Tat siRNA by our 4M5.3C construct is able to directly inhibit HIV infection.

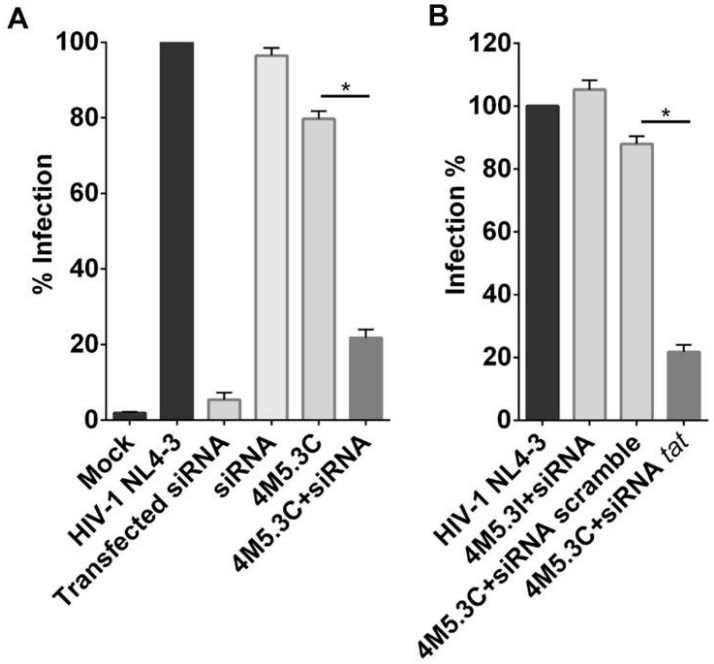


Figure 3.4: 4M5.3C-mediated delivery of tat siRNA inhibits HIV replication.
A) Percentage of HIV inhibition of T-lymphocytic SupT1 cell line infected with NL4-3 strain followed by treatment with 4M5.3C-siRNA anti-tat conjugation. After one day of NL4-3 infection, SupT1 cells were incubated in the presence of tat siRNA (siRNA) or 4M5.3C alone, or 4M5.3C+siRNA conjugation. "Mock" represents the non-infected cells. "HIV-1 NL4-3" represents infected cells with HIV-1 NL4-3 strain. "Transfected siRNA" represents lipid-based siRNA transfection for positive control of HIV inhibition. **B)** Percentage of HIV inhibition of NL4-3 infected SupT1 cells treated with CXCR4-targeted 4M5.3C construct conjugated with tat siRNA (4M5.3C+siRNA scramble) or scramble siRNA (4M5.3C+siRNA scramble), or irrelevant 4M5.3I control conjugated with tat siRNA (4M5.3I+siRNA tat). HIV replication was measured by viral p24 capsid quantification 72 hours after treatment. Values represent mean ± SEM of 2 independent assays with duplicates and are normalized to HIV-1 NL4-3. 4M5.3C+siRNA tat vs 4M5.3I+siRNA or vs 4M5.3C+siRNA scramble*$p<0.05$ (Mann-Whitney test).

3.4.4 CXCR4-mediated endocytosis of engineered zinc-finger transcription factor for HIV repression

To further explore NbCXCR4 capacity to potentiate delivery of HIV inhibitors, we developed an alternative to siRNA vehicle by directly fusing this variable domain with a zinc-finger transcription factor (ZF-TF) for HIV repression. Theoretically, any DNA sequence can be targeted by zinc-finger proteins that associated with functional domains such as transcription activator or repressors, are able to modulate gene expression. Even though zinc-binding domains were described as cell-penetrating proteins,[323] the absence of a receptor-directed delivery system compromises the efficiency of these domains, increasing the dosage and potentiating off-target effects in non-relevant populations.

Accordingly, we substituted the scFv portion of our delivery chimera by a ZF-TF (KRAB-HLTR3) towards the LTR HIV promoter, originating the KRAB-HLTR3C fusion protein (Fig. 3.5A). Among several ZF-TFs described to abrogate HIV replication[332,333,336,337] KRAB-HLTR3 is one of the most potent presenting 100-fold HIV repression and target conservation among HIV-1 clades, mainly clade B[333]. HLTR3

constitutes the zinc-finger domain, whereas the KRAB is the transcriptional repressor moiety. Inhibitory capacity of KRAB-HLTR3 is related with affinity and chromatin-accessibility of HLTR3 zinc-finger towards its target, repression potency of KRAB domain and the fact that HLTR3 binding-site overlaps the binding site of endogenous SP1 transcription factors—required for HIV activation[333]. We also constructed a non-targeting control with nanobody irrelevant as targeting moiety, KRAB-HLTR3I (Fig. 3.5A). The two fusion proteins as well as the NbCXCR4 and KRAB-HLTR3 alone were purified from *E. coli*, being the purity confirmed by SDS-PAGE (Fig. 3.5B).

We compared the entry efficiency between CXCR4-targeted KRAB-HLTR3 and zinc-finger repressor alone (KRAB-HLTR3) or in conjugation with the irrelevant nanobody (KRAB-HLR3I). Cellular extracts of $CXCR4^+$ T-cells incubated with the diverse ZF-TF constructions were analyzed by Western-blot for detection of HA-tagged constructs. As shown in Fig. 3.5C, there are no apparent differences on ZF-TFs uptake among the distinct constructions. However, when receptor-mediated endocytosis occurs (KRAB-HLTR3C) the nanobody portion of the chimera alone is detected (~16 kDa) along with the entire nanobody-ZF-TF conjugation (~49 kDa). In contrast, when ZF-TF is taken up into cells through non-specific ZF cell-penetration, KRAB-HLTR3 and KRAB-HLTR3I, only the whole zinc-finger repressor in fusion with irrelevant nanobody (~50 kDa) or alone (~35 kDa) are detected by western-blot. We hypothesized that the linker region between the ZF-TF and nanobody portions is particularly vulnerable to the harsh conditions of endosomal environment, leading to the liberation of the zinc-finger repressor and nanobody molecules as individual entities into cell cytoplasm. The observed KRAB-HLTR3C chimera within cells may enter by non-specific ZF route instead of CXCR4 endocytosis. To evaluate functionality of CXCR4-endocyted ZF-TF, we assessed repression of HIV LTR promoter into $CXCR4^+$ HeLaTat-III/LTR/d1EGFP cell line treated with KRAB-HLTR3C. This cell line encodes the HIV LTR promoter driving expression of a destabilized variant 1 of EGFP reporter (d1EGFP) and viral *tat* transactivator, which in turn causes a positive feedback on the LTR promoter to further enhance d1EGFP expression. Following protein treatment, cells were stained with labelled Alexa405-annexin V to exclude apoptotic cells and eliminate false LTR shutdown from non-viable cells. The KRAB-HLTR3 alone provided the positive control for repression of viral LTR promoter. Surprisingly, only the KRAB-HLTR3 in fusion with NbCXCR4 reduced d1EGFP fluorescence intensity of HeLa reporter line, indicating

ZF-TF mediated shutdown of LTR promoter (Fig. 3.5D). Non-targeting KRAB-HLTR3I control or NbCXCR4 alone had no effect on gene transcription from LTR promoter. Also, CXCR4 internalization due to nanobody-induced endocytosis has no negative effect on gene transcription from LTR promoter as demonstrated by the control of NbCXCR4 alone. Our results demonstrate that KRAB-HLTR3C is able to target and silence HIV promoter in a targeted manner through CXCR4 endocytosis.

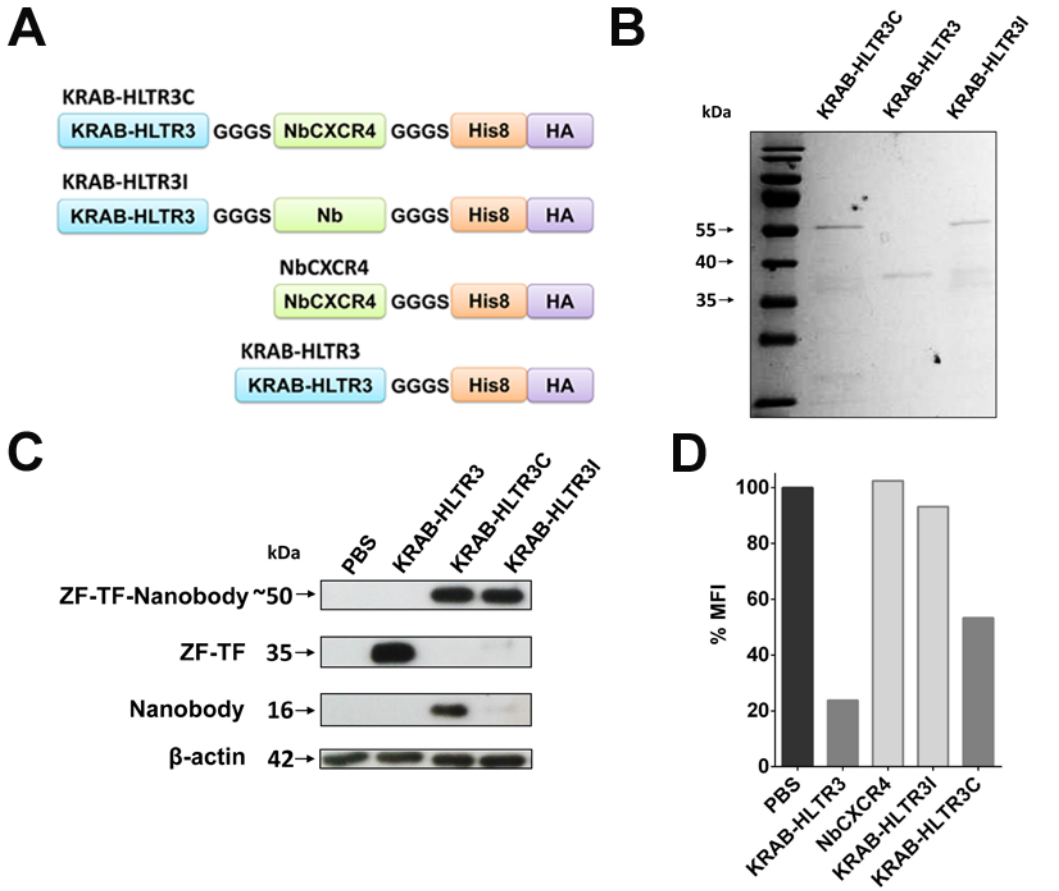


Figure 3.5: Delivery of zinc-finger transcription factor for repression of HIV promoter through CXCR4-targeting nanobody.
A) Schematic representation of ZF-TF delivery constructs. KRAB-HLTR3 zinc-finger repressor is positioned at the N-terminal of CXCR4-targeted (NbCXCR4) or irrelevant (Nb) nanobody to generate KRAB-HLTR3C or KRAB-HLTR3I, respectively. NbCXCR4 control construct is devoid of HLTR3-KRAB repressor domain, while KRAB-HLTR3 control is devoid of NbCXCR4 targeting domain. GGGS linkers were placed between the nanobody portion and the KRAB-HLTR3 domain or histidine (His) and hemagglutinin A (HA) tags, respectively for protein purification or detection**. B)** SDS-PAGE analysis of purified protein constructs. Proteins were detected by Coomassie-blue staining. **C)** Western blot of lysate of CXCR4+ Jurkat T-cells treated with KRAB-HLTR3 alone or in fusion with anti-CXCR4 (KRAB-HLTR3C) or irrelevant nanobodies (KRAB-HLTR3I). After 3 h of protein treatment, cells were treated with trypsin to eliminate the surface-bound proteins and the extracts analyzed by western-blot. Proteins were detected by peroxidase-conjugated anti-HA tag antibody, using anti-β-actin for loading control. "PBS" indicates lysate from untreated cells. **D)** Percentage of normalized d1EGFP Mean Fluorescence Intensity (MFI) of HeLa-Tat-III/LTR/d1EGFP treated with KRAB-HLTR3 or NbCXCR4 alone, or KRAB-HLTR3 in fusion with anti-CXCR4 (KRAB-HLTR3C) or irrelevant nanobodies (KRAB-HLTR3I). "PBS" indicates untreated cells. Cells were analyzed by flow cytometry 3 h after protein treatment for detection of d1EGP. Flow cytometry analysis of d1EGFP expression was gated on live cells by excluding apoptotic cells stained with Alexa 405-annexin V. d1EGFP MFI is normalized to untreated PBS control gated on viable cells negative for Alexa 405.

3.5 Discussion

Potency and breath of gene silencing methodologies hold promise for treatment of numerous diseases. Nevertheless, the lack of efficient and specific delivery methods continues to postpone the application of these technologies for human therapeutics.

Despite its efficacy for treatment of solid tumors,[411] delivery of therapeutic siRNAs in non-cancer applications is limited to disorders with simplified target-site accessibility such as ocular conditions or propensity to uptake/accumulate circulating molecules, as occurs with liver-located diseases.[411,412] This is not the case for HIV infection. Despite remarkable advances in current antiretroviral therapy, HIV infection remains as a life-threatening condition with serious morbidity effects. Accordingly, development of versatile HIV expression inhibitors such as RNAi,[261] or additionally zinc-finger repressors,[332,333,336,337] could provide a volte-face in this viral disorder therapeutics. However, inhibitors targeting HIV expression require a systemic delivery system due to the widespread distribution of virus-susceptible cells through human body.

Here, we present targeted approaches to deliver inhibitors of HIV expression (siRNA and zinc-finger repressors) through engineered nanobody-based chimeras. Exceptional characteristics of nanobodies including high stability and solubility, low immunogenicity and increased tissue penetration[25] make their a suitable option for targeting HIV-infected cells primarily located onto densely packed tissues. The absence of an effector region (Fc) in these antibody formats constitutes an additional safety feature to prevent undesirable immune activation. We developed a novel method for specific delivery of RNAi effectors, using a scFv-nanobody chimera (4M5.3C) for CXCR4-targeting of an anti-HIV siRNA. This chimera conjugates a validated nanobody 238D2[406] responsible for membrane translocation through CXCR4-mediated endocytosis with the 4M5.3 antibody fragment, a scFv with exquisitely binding affinity towards FITC fluorochrome[405] for promoting efficient uptake of FITC-conjugated siRNA effectors.

In this study, we demonstrate 4M5.3C-mediated delivery of anti-*tat* siRNA to CXCR4-bearing cells, abolishing transcription from HIV LTR promoter. These data confirm the specificity of our delivery strategy towards $CXCR4^+$ cells and ensures the silencing functionality of the delivered therapeutic siRNA. Delivery of 4M5.3C+siRNA conjugates also impaired viral replication in a T-lymphocytic cell line. Overall, these results validate the reported 4M5.3C chimera as a vehicle for delivery of anti-HIV siRNAs further demonstrating the potential of ligand-dependent CXCR4 endocytosis as

an entry route for therapeutic molecules as others already described.[397–399,413] Few others reported targeted delivery of anti-HIV siRNAs, specially through antibody-based vehicles.[261] Moreover, none described a nanobody format as targeting moiety.[250,251,414]

In this chapter, we also fused a zinc-finger repressor to the NbCXCR4 (KRAB-HLTR3C) for further testing a strategy of endocytosis-dependent functionality. Zinc-finger domains have an intrinsic cell permeability,[323] which surpasses the need for a delivery system. However, ZFPs lack of specificity leads to an inefficient therapeutic scheme and indesirable off-target effects. Opposed to non-targeted irrelevant chimera (4M5.3I), we detected the NbCXCR4 portion disproved of HLTR3-KRAB repressor into the intracellular environment, which indicates chimera hydrolysis between the nanobody and transcription factor portions in consequence of endocytosis delivery. On the other hand, only the CXCR4-targeted KRAB-HLTR3 exhibited repression of gene expression from HIV LTR promoter. These data seem to indicate that targeted-receptor endocytosis is required to assure zinc-finger repressor release and access to the cellular nucleus to perform gene modulation. Moreover, we did not observe any significant cytotoxicity associated with our chimera treatment. We believe that this should be related with the low range of protein concentration (not surpassing 0.2 μM) required for therapeutic benefit, derived from the increased potency of our strategy for enhanced ZF-TF delivery mediated by CXCR4-endocytosis. Previously, Meiler and co-workers already described a ligand-mediated delivery of ZFPs through transferrin receptor for ZF-based nucleases,[415] demonstrating that endocytosis-mediated delivery enhanced ZF cell-penetration and consequently its activity. Other study had also reported successfully delivery of zinc-finger repressors, although through a non-specific system of protein transduction domain.[416] Overall, these results support the potential of targeted delivery of KRAB-HLTR3 repressor to inhibit HIV replication. Further studies have now to be performed to confirm this hypothesis such as prove of KRAB-HLTR3 presence on the nucleus only when delivered through CXCR4 endocytosis.

We anticipate that our strategy for CXCR4-targeted delivery of siRNA or zinc-finger inhibitors will provide broad inhibition of HIV replication apart from its tropism, as all infected cells possess this coreceptor.[417,418] The usage of a cellular receptor for targeted delivery also constitutes an advantage of our strategy relative to others,[248,250,419] since viral targets are prone to mutational modifications and limit therapeutic effect to already infected and actively replicating (non-latent) cells. Also, we cannot exclude the

possibility of using this system as a dual-inhibitory approach for CXCR4-tropic virus, as nanobody portion of our chimera was previously validated as a fusion inhibitor of HIV infection through binding competition with viral co-receptor CXCR4.[406]

In conclusion, the presented study provides a new paradigm for delivery of HIV repressor molecules by CXCR4-targeted nanobody. Moreover, we expect that this method could pave the way for novel delivery systems of HIV inhibitors.

3.6 Acknowledgments

We thank D. Guo for kindly providing lentiCXCR4-gRNA-Cas9 #6, psPAX2 and pMD2.G plasmids, M. Simurda for kindly providing pTat plasmid, C. Rodrigues for kindly providing primary mouse anti-β-actin and C. Barbas III for kindly providing KRAB-HLTR3 plasmid.

C.C-S. and J.G. conceived and designed the experiments. C.C-S. performed the experiments, analysed the data and wrote the article.

This work was supported by grants HIVERA/0002/2013, UTAP-ICDT/DTP-FTO/0016/2014, VIH/SAU/0013/2011 and VIH/SAU/0020/2011 from Fundação para a Ciência e a Tecnologia—Ministério da Educação e Ciência (FCT-MEC), Portugal. CC-S acknowledge FCT-MEC for PhD fellowship SFRH/BD/73838/2010.

3.7 Supplementary Information

3.7.1 Supplementary Tables

Table 3.1: PCR fragments and primers used in this study.

scFv 4M5.3C αFITC for chimera construction

>scFv4M5.3-αFITC-F-NheI

5’ – GCACCACT AGCTAGCTACCCCT ACGACGTGCCTGATT ACGCCGGAGGAT CCGGAGGA - 3’

>scFv4M5.3-αFITC-R-SacI

5’ -AGTCTGGGTCAG CAGGAGCTCGCT TCCGCCTCCTCCGCT CACTGTCAC-3’

Nanobody 282D2 αCXCR4/Nanobody irrelevant for chimera construction

> Nanobody-F-SacI

5’-CGAGCT CGAGGTGC AGCTGG TGGA-3’

> Nanobody-R-XhoI

5’- CCGCTCGAGT CATGCGTAATC AGGCACGTCGTA GGGGTACGATCC ATGGTGATG GTGATGGTGATG GTGGCTGCCTCCGC CTCCACTGCTGCTCAC GGTCACCTG -3’

Nanobody 282D2 αCXCR4 alone

> Nanobody-F-NheI

5’-CTAGCT AGCGAGGTGC AGCTGGTGGA-3’

> Nanobody-R-XhoI

5’- CCGCT CGAGTCATGCGT AATCAGGCACGTCGTA GGGGTACGAT CCATGG TGATGGTGATG GTGATGGTGGCT GCCTCCGCCTCC ACTGCTGCTCA CGGTCACCTG -3’

KRAB-HLTR3 for chimera construction

> KRAB-HLTR3-F-NheI

5’- GGAATTCCATATGGATGCTAAGTCACTGACTGCCT -3’

> KRAB-HLTR3-R-SacI

5’- CGAGCTCGCTTCCGCCTCCGCTACTAGTTTTTTTACC -3’

KRAB-HLTR3 alone

> KRAB-HLTR3-F-NheI

5’- GGAATTCCATATGGATGCTAAGTCACTGACTGCCT -3’

> KRAB-HLTR3-R-XhoI

5’- CCGCTCGAGTCAT GCGTAATCA GGCACG TCGTAGGGG TACGATCC ATGGTGATGG TGATGGTGATG GTGGCTGCCTCCGCCTC CGCTACTAGTTTTTTTACC -3’

CHAPTER IV

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

Recent advances on antibody engineering towards novel scaffolds and formats have paved the way for an innovative and promising class of biotherapeutic molecules. The antigen-binding sites of single-domain antibody (sdAb) fragments are particularly suitable to be tailored to create high affinity and specific binders. On the other hand, the peculiar heavy-chain only antibodies from camels allows to isolate autonomous variable domains (nanobodies) with favorable physical characteristics and exquisite binding capacity towards non-standard epitopes.

Despite progresses on cART development and implementation, HIV infection continues to be a major worldwide hurdle due to disease persistence, viral resistance, and complications/toxicity of standard therapy. The exploration of innovative therapeutic approaches together with the increasing knowledge on virus pathogenesis could bring a new era of antiretroviral molecules for control of this lifetime disease. Accordingly, promising classes of HIV inhibitors are now beginning to be explored for clinical purposes such as antibody-, RNAi- and ZF-based molecules (aidsinfo.nih.gov). However, the lack of efficient and specific delivery systems could constitute a serious drawback on the implementation of these effector molecules for human therapy as antiviral drugs. Overall, there is urgent need to design and engineer more potent and broad therapeutic approaches for HIV infection as well as the platforms to their efficient and specific delivery.

The present thesis set out to explore the potential of engineered sdAbs for treatment of HIV/AIDS condition, either as broad and potent inhibitors of HIV infection or as vehicles for targeted antiretroviral delivery.

In Chapter I, we presented a general introduction to the antibody engineering field, in particular single-domain antibodies characteristics and state-of-the-art in therapeutic applications as well as synthetic libraries tailoring. We also introduced HIV virus biology and therapeutics with focus on unmet needs and drawbacks of available treatment scheme.

In Chapter II, we have explored a strategy for the design of synthetic enlarged CDRs to develop a sdAb inhibitor of HIV infection. For this purpose, we took advantage of a previously validated rabbit scaffold to construct a library of light-chain antibodies (VL domains) with combinatorial restriction bias natural present residues at antigenic sites. This approach pretended to assure library's quality through high percentage of functional clones and intended high-affinity binders. VL fragments were selected by

phage display technology for direct binding to the crucial-to-fusion and conserved N36 region on viral envelope glycoprotein. The improved penetrability of sdAbs fragments make them more suitable to target sterically occluded cavities than alternative formats such as Fabs and scFvs. Because of extremely convex nature of N36 region, we designed sdAbs harboring long and flexible CDRs to further improve target accessibility by these fragments.[119,420–422] We narrowed five potential HIV inhibitors with specific binding to N36 peptide to one VL showing broad HIV neutralization, herein named F63. The potency and breath activity of F63, evidenced through the proved neutralization of distinct HIV-1 and HIV-2 strains, also confirms the conservation and importance of N36 region for virus infection as well as rationalize its immune evasion nature. On the other hand, the verified F63 propensity for lipid membranes interaction is shared with other potent entry inhibitors of HIV infection.[346,347] Nonetheless, further evaluation of *in vivo* efficacy/safety of F63 is required.

Our data seem to validate the selection of long CDRs bias cleft epitopes, evidencing the successfully translation of protuberating nature from camel and shark CDRs into alternative scaffolds. Additionally, our results reveal good perspectives for the screening of this synthetic library against alternative non-standard (cryptic) targets such as active sites of enzymes. Accordingly, this repertoire applicability could be exponentially expanded towards therapeutics of other disorder types or diagnostic/investigational approaches. The applicability of this dAb does not also have to be limited to neutralization of cell-free virus and inhibition of cell-cell HIV transmission. This fusion inhibitor can be further tested as microbicide or PrEP (pre-exposure prophylaxis) drug for prevention of HIV acquisition. Also, it would be interesting to evaluate dAb capacity to deliver toxic payloads to HIV-bearing cells, setting this minimal antibody fragment with a dual-antiviral activity, rare in HIV inhibitors. This potential is supported by previous studies describing antibodies targeting epitopes on gp41 HR loops as the most effective immunotoxins (antibody-toxin conjugations).[423,424] In resume, the data presented in Chapter II evidence the essential role of CDRs engineering, either in diversity or length, for the successfully design of synthetic libraries, amenable to be screened against therapeutically relevant and difficult-to-access targets such as cryptic N36 region on HIV surface glycoprotein. Additionally, these results provide proof-of-principle that engineered rabbit domain antibodies exhibit remarkable efficiency and cross-reacting neutralization of HIV-1 and HIV-2 infection, establishing a novel class of

potent viral inhibitors based on light-chain sdAbs. These domains can also help to identify original conserved epitopes as well as elucidate the virus mechanisms of entry and evasion from immune system.

In Chapter III, we developed a strategy for specific delivery of novel classes of HIV inhibitors, in this case repressors of viral gene expression. We took advantage of a previously validated nanobody towards CXCR4 receptor to engineer a fusion construction for targeted delivery of an anti-HIV small interfering RNA (siRNA). This strategy is based on the design of a nanobody targeting domain fused with an anti-FITC scFv fragment to carry FITC-conjugated siRNA. The nanobody-based chimera (4M5.3C) delivered the anti-*tat* siRNA specifically to $CXCR4^+$ cells, being the functionality of the delivered RNAi inhibitor proved through silencing of gene transcription from Tat-driven virus promoter and prevention of HIV-1 replication. In addition, we constructed a chimera for targeted delivery of an engineered zinc-finger repressor of HIV (KRAB-HLTR3C) through CXCR4 endocytosis. Incorporation of a nanobody fused to zinc-finger repressor turned its activity dependent of endocytosis-mediated entry, with nanobody-fused zinc-finger chimera presenting shutdown of HIV LTR promoter only when targeted to CXCR4 receptor. Our results indicate the release of zinc-finger repressor from nanobody chimera, probably in consequence of endosomal harsh conditions following CXCR4-endocytosis, and a tendency for the hydrolysis of the linker sequence between the zinc-finger repressor and the nanobody portions.

The results from this Chapter validate sdAbs applicability for specific delivery of antiviral inhibitors, evidencing the efficacy of the targeted molecules as repressors of HIV gene expression. Despite the reported positive results, *in vivo* studies must validate this nanobody-based strategy for systemic delivery of anti-HIV siRNAs and zinc-finger repressors through CXCR4-mediated endocytosis. Furthermore, this strategy may present bias towards virus-infected cells. This phenomenon is supported by HIV requirement and perpetuation in activated T-cells, where CXCR4 expression is up-regulated.[425] It would be interesting to evaluate capacity of scFv-nanobody construct to provide a shield against serum nucleases-mediated degradation of siRNA although chemical modifications on carried siRNA may also provide the same effect.[261] The ability of scFv-nanobody chimera to mask therapeutic siRNA from immune system would be also of value to explore. Extension of siRNA half-life should be confirmed by pharmacokinetics studies though

the size of the 4M5.3C+siRNA complex is estimated in ~52 kDa, well above the renal filtration cut-off (30 kDa). Both delivery strategies can also be tested against latently infected cells. Impairment of provirus (integrated HIV genome) activation is not only important to avoid virus spread as assumes a crucial role on "lock out" strategies for latent infection inactivation.[426] In these approaches, the full suppression of HIV provirus into a deep dormant status is expected to conduct to the natural elimination of HIV latent cells through cells renewal process. Our data anticipate a broad applicability of this strategy towards non-HIV diseases and within alternative therapeutic molecules. Alternative nanobody portions may modify delivery specificity of the chimera, as scFv fragment could also be replaced by RNA-binding molecules. Moreover, the possibility to multiplex siRNA molecules against diverse targets could enable therapeutically managing of distinct diseases within a single individual. For example, CXCR4 is an overexpressed receptor on malignant cells, constituting an attractive target for treatment of cancer diseases[406] such as AIDS-related lymphoma—common secondary condition of HIV-infected patients. Alternative zinc-finger effectors,[292] including site-directed nucleases, recombinases or transcription activators, are as well expected to be delivered by this strategy. In resume, these two strategies for HIV inhibitors delivery pretend to further demonstrate the potential of minimal antibody fragments for therapeutic purposes as component vehicles of antiretroviral effectors.

In conclusion, the present thesis provides proof-of-concept on single-domain antibodies usefulness in distinct therapeutic approaches for AIDS epidemic. We demonstrate the potential of rational design in tailoring a fit-to-purpose antibody fragment, in this case a broad fusion inhibitor of HIV infection based on a VL domain from rabbit-origin. On the other hand, we explored a nanobody-based platform for targeted delivery of distinctive technologies for HIV expression downregulation, either through zinc-finger transcriptional repressors or siRNA gene silencers. Accordingly, our study give insights into the engineering of highly attractive minimal domains for antiviral therapeutics, providing the basis for development of next-generation sdAbs for HIV/AIDS treatment.

CHAPTER V

REFERENCES

REFERENCES

1. Reichert, J. M. Antibodies to watch in 2017. *MAbs* 1–15 (2016). doi:10.1080/19420862.2016.1269580
2. Schroeder, H. W., Cavacini, L. & Cavacini, L. Structure and function of immunoglobulins. *J. Allergy Clin. Immunol.* **125,** S41-52 (2010).
3. Williams, A. F. & Barclay, A. N. The immunoglobulin superfamily--domains for cell surface recognition. *Annu. Rev. Immunol.* **6,** 381–405 (1988).
4. Roopenian, D. C. & Akilesh, S. FcRn: the neonatal Fc receptor comes of age. *Nat. Rev. Immunol.* **7,** 715–25 (2007).
5. Brekke, O. H. & Sandlie, I. Therapeutic antibodies for human diseases at the dawn of the twenty-first century. *Nat. Rev. Drug Discov.* **2,** 52–62 (2003).
6. Chan, A. C. & Carter, P. J. Therapeutic antibodies for autoimmunity and inflammation. *Nat. Rev. Immunol.* **10,** 301–316 (2010).
7. MacGlashan, D. W. Endocytosis, recycling, and degradation of unoccupied FcepsilonRI in human basophils. *J. Leukoc. Biol.* **82,** 1003–10 (2007).
8. Köhler, G. & Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. 1975. *J. Immunol.* **174,** 2453–2455 (2005).
9. Khazaeli, M. B., Conry, R. M. & LoBuglio, A. F. Human immune response to monoclonal antibodies. *J. Immunother. Emphasis Tumor Immunol.* **15,** 42–52 (1994).
10. Jones, P., Dear, P., Foote, J., Neuberger, M. & Winter, G. Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature* **321,** 522–5 (1986).
11. Morrison, S. L., Johnson, M. J., Herzenberg, L. a & Oi, V. T. Chimeric human antibody molecules: mouse antigen-binding domains with human constant region domains. *Proc. Natl. Acad. Sci. U. S. A.* **81,** 6851–6855 (1984).
12. Clackson, T., Hoogenboom, H. R., Griffiths, A. D. & Winter, G. Making antibody fragments using phage display libraries. *Nature* **352,** 624–628 (1991).
13. Green, L. L. *et al.* Antigen-specific human monoclonal antibodies from mice engineered with human Ig heavy and light chain YACs. *Nat. Genet.* **7,** 13–21 (1994).
14. Bird, R. *et al.* Single-chain antigen-binding proteins. *Science (80-.).* **242,** 423–426 (1988).
15. Ward, E. S., Güssow, D., Griffiths, a D., Jones, P. T. & Winter, G. Binding activities of a repertoire of single immunoglobulin variable domains secreted from Escherichia coli. *Nature* **341,** 544–546 (1989).
16. Hamers-Casterman, C. *et al.* Naturally occurring antibodies devoid of light chains. *Nature* **363,** 446–8 (1993).
17. Holliger, P. & Hudson, P. J. Engineered antibody fragments and the rise of single domains. *Nat. Biotechnol.* **23,** 1126–36 (2005).
18. Little, M. *Recombinant antibodies for immunotherapy*. (Cambridge University Press, 2009).
19. Wark, K. L. & Hudson, P. J. Latest technologies for the enhancement of antibody affinity. *Adv. Drug Deliv. Rev.* **58,** 657–70 (2006).
20. Kontermann, R. E. Strategies for extended serum half-life of protein therapeutics. *Curr. Opin. Biotechnol.* **22,** 868–876 (2011).

REFERENCES

21. Adolf-Bryfogle, J., Xu, Q., North, B., Lehmann, A. & Dunbrack, R. L. PyIgClassify: a database of antibody CDR structural classifications. *Nucleic Acids Res.* **43,** D432–D438 (2015).
22. Collis, A. V. J., Brouwer, A. P. & Martin, A. C. R. Analysis of the antigen combining site: Correlations between length and sequence composition of the hypervariable loops and the nature of the antigen. *J. Mol. Biol.* **325,** 337–354 (2003).
23. Sela-Culang, I., Kunik, V. & Ofran, Y. The structural basis of antibody-antigen recognition. *Front. Immunol.* **4,** 302 (2013).
24. Lavinder, J. J., Hoi, K. H., Reddy, S. T., Wine, Y. & Georgiou, G. Systematic characterization and comparative analysis of the rabbit immunoglobulin repertoire. *PLoS One* **9,** e101322 (2014).
25. Hamers-Casterman, C. *et al.* Naturally occurring antibodies devoid of light chains. *Nature* **363,** 446–448 (1993).
26. Greenberg, A. S. *et al.* A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. *Nature* **374,** 168–73 (1995).
27. Alexander, A. *et al.* gamma Heavy chain disease in man: cDNA sequence supports partial gene deletion model. *Proc. Natl. Acad. Sci. U. S. A.* **79,** 3260–4 (1982).
28. Cogné, M., Preud'homme, J. L. & Guglielmi, P. Immunoglobulin gene alterations in human heavy chain diseases. *Res. Immunol.* **140,** 487–502
29. Desmyter, A. *et al.* Three camelid VHH domains in complex with porcine pancreatic α-amylase: Inhibition and versatility of binding topology. *J. Biol. Chem.* **277,** 23645–23650 (2002).
30. Desmyter, A. *et al.* Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nat. Struct. Biol.* **3,** 803–11 (1996).
31. De Genst, E. *et al.* Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 4586–4591 (2006).
32. Muyldermans, S., Atarhouch, T., Saldanha, J., Barbosa, J. A. & Hamers, R. Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. *Protein Eng.* **7,** 1129–35 (1994).
33. Barthelemy, P. a. *et al.* Comprehensive analysis of the factors contributing to the stability and solubility of autonomous human VH domains. *J. Biol. Chem.* **283,** 3639–3654 (2008).
34. Barelle, C. & Porter, A. VNARs: An Ancient and Unique Repertoire of Molecules That Deliver Small, Soluble, Stable and High Affinity Binders of Proteins. *Antibodies* **4,** 240–258 (2015).
35. Stanfield, R. L., Dooley, H., Flajnik, M. F. & Wilson, I. A. Crystal structure of a shark single-domain antibody V region in complex with lysozyme. *Science* **305,** 1770–3 (2004).
36. Flajnik, M. F., Deschacht, N. & Muyldermans, S. A case of convergence: why did a simple alternative to canonical antibodies arise in sharks and camels? *PLoS Biol.* **9,** e1001120 (2011).
37. Zielonka, S. *et al.* Structural insights and biomedical potential of IgNAR scaffolds from sharks. *mAbs* **7,** 15–25 (2015).
38. Doshi, R. *et al.* In vitro nanobody discovery for integral membrane protein targets. *Sci. Rep.* **4,** 6760 (2014).
39. Riechmann, L. & Muyldermans, S. Single domain antibodies: comparison of camel VH and camelised human VH domains. *J. Immunol. Methods* **231,** 25–38 (1999).

40. Davies, J. & Riechmann, L. 'Camelising' human antibody fragments: NMR studies on VH domains. *FEBS Lett.* **339,** 285–290 (1994).
41. Tanha, J. *et al.* Optimal design features of camelized human single-domain antibody libraries. *J. Biol. Chem.* **276,** 24774–80 (2001).
42. Davies, J. & Riechmann, L. Single antibody domains as small recognition units: design and in vitro antigen selection of camelized, human VH domains with improved protein stability. *Protein Eng.* **9,** 531–537 (1996).
43. Da Silva, F. A. *et al.* Camelized rabbit-derived VH single-domain intrabodies against Vif strongly neutralize HIV-1 infectivity. *J. Mol. Biol.* **340,** 525–542 (2004).
44. Dottorini, T., Vaughan, C. K., Walsh, M. A., LoSurdo, P. & Sollazzo, M. Crystal structure of a human VH: requirements for maintaining a monomeric fragment. *Biochemistry* **43,** 622–8 (2004).
45. Ewert, S., Cambillau, C., Conrath, K. & Plückthun, A. Biophysical properties of camelid V(HH) domains compared to those of human V(H)3 domains. *Biochemistry* **41,** 3628–36 (2002).
46. Colby, D. W. *et al.* Potent inhibition of huntingtin aggregation and cytotoxicity by a disulfide bond-free single-domain intracellular antibody. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 17616–21 (2004).
47. Hussack, G. *et al.* A VL single-domain antibody library shows a high-propensity to yield non-aggregating binders. *Protein Eng. Des. Sel.* **25,** 313–318 (2012).
48. Ewert, S., Huber, T., Honegger, A. & Plückthun, A. Biophysical properties of human antibody variable domains. *J. Mol. Biol.* **325,** 531–553 (2003).
49. Hussack, G., Hirama, T., Ding, W., MacKenzie, R. & Tanha, J. Engineered single-domain antibodies with high protease resistance and thermal stability. *PLoS One* **6,** (2011).
50. Chan, P.-H. *et al.* Engineering a camelid antibody fragment that binds to the active site of human lysozyme and inhibits its conversion into amyloid fibrils. *Biochemistry* **47,** 11041–54 (2008).
51. Hagihara, Y., Mine, S. & Uegaki, K. Stabilization of an immunoglobulin fold domain by an engineered disulfide bond at the buried hydrophobic region. *J. Biol. Chem.* **282,** 36489–95 (2007).
52. Kim, D. Y. *et al.* Disulfide linkage engineering for improving biophysical properties of human VH domains. *Protein Eng. Des. Sel.* **25,** 581–9 (2012).
53. Kim, D. Y., Ding, W. & Tanha, J. Solubility and stability engineering of human VH domains. *Methods Mol. Biol.* **911,** 355–72 (2012).
54. Saerens, D., Conrath, K., Govaert, J. & Muyldermans, S. Disulfide bond introduction for general stabilization of immunoglobulin heavy-chain variable domains. *J. Mol. Biol.* **377,** 478–88 (2008).
55. Ma, X., Barthelemy, P. A., Rouge, L., Wiesmann, C. & Sidhu, S. S. Design of synthetic autonomous VH domain libraries and structural analysis of a VH domain bound to vascular endothelial growth factor. *J. Mol. Biol.* **425,** 2247–59 (2013).
56. Rouet, R., Dudgeon, K., Christie, M., Langley, D. & Christ, D. Fully Human VH Single Domains That Rival the Stability and Cleft Recognition of Camelid Antibodies. *J. Biol. Chem.* **290,** 11905–17 (2015).
57. Cai, X. & Garen, A. A melanoma-specific VH antibody cloned from a fusion phage library of a vaccinated melanoma patient. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 6280–5 (1996).
58. Jespers, L., Schon, O., Famm, K. & Winter, G. Aggregation-resistant domain antibodies selected

on phage by heat denaturation. *Nat. Biotechnol.* **22,** 1161–5 (2004).

59. Tanha, J. *et al.* Improving solubility and refolding efficiency of human V(H)s by a novel mutational approach. *Protein Eng. Des. Sel.* **19,** 503–9 (2006).
60. Jespers, L., Schon, O., James, L. C., Veprintsev, D. & Winter, G. Crystal structure of HEL4, a soluble, refoldable human V(H) single domain with a germ-line scaffold. *J. Mol. Biol.* **337,** 893–903 (2004).
61. To, R. *et al.* Isolation of monomeric human V(H)s by a phage selection. *J. Biol. Chem.* **280,** 41395–403 (2005).
62. Famm, K., Hansen, L., Christ, D. & Winter, G. Thermodynamically stable aggregation-resistant antibody domains through directed evolution. *J. Mol. Biol.* **376,** 926–31 (2008).
63. Dubnovitsky, a P. *et al.* Expression, refolding, and ferritin-binding activity of the isolated VL-domain of monoclonal antibody F11. *Biochem. Biokhimiia* **65,** 1011–8 (2000).
64. Schiefner, A. *et al.* A Disulfide-Free Single-Domain VL Intrabody with Blocking Activity towards Huntingtin Reveals a Novel Mode of Epitope Recognition. *J. Mol. Biol.* **414,** 337–355 (2011).
65. Lee, W.-R., Jang, J.-Y., Kim, J.-S., Kwon, M.-H. & Kim, Y.-S. Gene silencing by cell-penetrating, sequence-selective and nucleic-acid hydrolyzing antibodies. *Nucleic Acids Res.* **38,** 1596–1609 (2009).
66. Paz, K. *et al.* Human single-domain neutralizing intrabodies directed against Etk kinase: a novel approach to impair cellular transformation. *Mol. Cancer Ther.* **4,** 1801–9 (2005).
67. Colby, D. W. *et al.* Development of a human light chain variable domain (V L) intracellular antibody specific for the amino terminus of huntingtin via yeast surface display. *J. Mol. Biol.* **342,** 901–912 (2004).
68. van den Beucken, T. *et al.* Building novel binding ligands to B7.1 and B7.2 based on human antibody single variable light chain domains. *J. Mol. Biol.* **310,** 591–601 (2001).
69. Kim, D. Y. *et al.* Antibody light chain variable domains and their biophysically improved versions for human immunotherapy. *MAbs* **6,** 219–235 (2014).
70. Brinkmann, U., Lee, B. K. & Pastan, I. Recombinant immunotoxins containing the VH or VL domain of monoclonal antibody B3 fused to Pseudomonas exotoxin. *J. Immunol.* **150,** 2774–82 (1993).
71. Galan, A. *et al.* Library- based display technologies: where do we stand? *Mol. BioSyst.* 2342–2358 (2016). doi:10.1039/C6MB00219F
72. Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228,** 1315–7 (1985).
73. Sidhu, S. S. & Koide, S. Phage display for engineering and analyzing protein interaction interfaces. *Curr. Opin. Struct. Biol.* **17,** 481–7 (2007).
74. Feeney, A. J. Genetic and epigenetic control of V gene rearrangement frequency. *Adv. Exp. Med. Biol.* **650,** 73–81 (2009).
75. Li, A. *et al.* Utilization of Ig heavy chain variable, diversity, and joining gene segments in children with B-lineage acute lymphoblastic leukemia: implications for the mechanisms of VDJ recombination and for pathogenesis. *Blood* **103,** 4602–9 (2004).

76. Mauerer, K. *et al.* Immunoglobulin gene segment usage, location and immunogenicity in mutated and unmutated chronic lymphocytic leukaemia. *Br. J. Haematol.* **129,** 499–510 (2005).
77. Miersch, S. & Sidhu, S. S. Synthetic antibodies: Concepts, potential and practical considerations. *Methods* **57,** 486–498 (2012).
78. Neylon, C. Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nucleic Acids Res.* **32,** 1448–59 (2004).
79. Fellouse, F. a, Wiesmann, C. & Sidhu, S. S. Synthetic antibodies from a four-amino-acid code: a dominant role for tyrosine in antigen recognition. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 12467–12472 (2004).
80. Fellouse, F. a. *et al.* High-throughput Generation of Synthetic Antibodies from Highly Functional Minimalist Phage-displayed Libraries. *J. Mol. Biol.* **373,** 924–940 (2007).
81. Fellouse, F. a *et al.* Molecular recognition by a binary code. *J. Mol. Biol.* **348,** 1153–62 (2005).
82. Birtalan, S. *et al.* The intrinsic contributions of tyrosine, serine, glycine and arginine to the affinity and specificity of antibodies. *J. Mol. Biol.* **377,** 1518–28 (2008).
83. Koide, S. & Sidhu, S. S. The importance of being tyrosine: lessons in molecular recognition from minimalist synthetic binding proteins. *ACS Chem. Biol.* **4,** 325–334 (2009).
84. Villar, H. O. & Kauvar, L. M. Amino acid preferences at protein binding sites. *FEBS Lett.* **349,** 125–30 (1994).
85. Zemlin, M. *et al.* Expressed murine and human CDR-H3 intervals of equal length exhibit distinct repertoires that differ in their amino acid composition and predicted range of structures. *J. Mol. Biol.* **334,** 733–749 (2003).
86. Barré-Sinoussi, F. *et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220,** 868–71 (1983).
87. Gallo, R. C. *et al.* Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* **224,** 500–3 (1984).
88. Clavel, F. *et al.* Isolation of a new human retrovirus from West African patients with AIDS. *Science* **233,** 343–346 (1986).
89. Campbell-Yesufu, O. T. & Gandhi, R. T. Update on human immunodeficiency virus (HIV)-2 infection. *Clin. Infect. Dis.* **52,** 780–7 (2011).
90. Simon, F. *et al.* Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat. Med.* **4,** 1032–7 (1998).
91. Plantier, J.-C. *et al.* A new human immunodeficiency virus derived from gorillas. *Nat. Med.* **15,** 871–872 (2009).
92. Robertson, D. L. *et al.* HIV-1 nomenclature proposal. *Science* **288,** 55–6 (2000).
93. Gao, F. *et al.* Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. *J. Virol.* **68,** 7433–47 (1994).
94. Yamaguchi, J. *et al.* HIV type 2 intergroup recombinant identified in Cameroon. *AIDS Res. Hum. Retroviruses* **24,** 86–91 (2008).
95. Ibe, S. *et al.* HIV-2 CRF01_AB: first circulating recombinant form of HIV-2. *J. Acquir. Immune*

Defic. Syndr. **54,** 241–7 (2010).

96. Li, G. *et al.* An integrated map of HIV genome-wide variation from a population perspective. *Retrovirology* **12,** 18 (2015).
97. Arthur, L. O. *et al.* Cellular proteins bound to immunodeficiency viruses: implications for pathogenesis and vaccines. *Science* **258,** 1935–8 (1992).
98. Taube, R., Fujinaga, K., Wimmer, J., Barboric, M. & Peterlin, B. M. Tat transactivation: a model for the regulation of eukaryotic transcriptional elongation. *Virology* **264,** 245–253 (1999).
99. Jeang, K. T., Xiao, H. & Rich, E. A. Multifaceted activities of the HIV-1 transactivator of transcription, Tat. *J. Biol. Chem.* **274,** 28837–28840 (1999).
100. He, G., Ylisastigui, L. & Margolis, D. M. The regulation of HIV-1 gene expression: the emerging role of chromatin. *DNA Cell Biol.* **21,** 697–705 (2002).
101. Shum, K. T., Zhou, J. & Rossi, J. J. Aptamer-based therapeutics: New approaches to combat human viral diseases. *Pharmaceuticals* **6,** 1507–1542 (2013).
102. Hedestam, G. B. K. *et al.* The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus. *Nat. Rev. Microbiol.* **6,** 143–155 (2008).
103. Peterlin, B. M. & Trono, D. Hide, shield and strike back: how HIV-infected cells avoid immune eradication. *Nat. Rev. Immunol.* **3,** 97–107 (2003).
104. Caputi, M. in *RNA Processing* (InTech, 2011). doi:10.5772/20899
105. Dalgleish, A. *et al.* The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* **312,** 763–767 (1984).
106. Klatzmann, D. *et al.* T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* **312,** 767–768 (1984).
107. Maddon, P. J. *et al.* The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* **47,** 333–348 (1986).
108. McDougal, J. S. *et al.* Binding of HTLV-III/LAV to T4+ T cells by a complex of the 110K viral protein and the T4 molecule. *Science* **231,** 382–385 (1986).
109. Liu, Y. *et al.* CD4-independent infection of astrocytes by human immunodeficiency virus type 1: requirement for the human mannose receptor. *J. Virol.* **78,** 4120–33 (2004).
110. Guttman, M., Kahn, M., Garcia, N. K., Hu, S.-L. & Lee, K. K. Solution Structure, Conformational Dynamics, and CD4-Induced Activation in Full-Length, Glycosylated, Monomeric HIV gp120. *J. Virol.* **86,** 8750–8764 (2012).
111. Diskin, R., Marcovecchio, P. M. & Bjorkman, P. J. Structure of a clade C HIV-1 gp120 bound to CD4 and CD4-induced antibody reveals anti-CD4 polyreactivity. *Nat. Struct. Mol. Biol.* **17,** 608–613 (2010).
112. Kwong, P. D. *et al.* HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. *Nature* **420,** (2002).
113. Berger, E. A. *et al.* A new classification for HIV-1. *Nature* **391,** 240 (1998).
114. Simmons, G. *et al.* Co-receptor use by HIV and inhibition of HIV infection by chemokine receptor ligands. *Immunol. Rev.* **177,** 112–26 (2000).
115. Zhang, L. *et al.* Chemokine coreceptor usage by diverse primary isolates of human

immunodeficiency virus type 1. *J. Virol.* **72,** 9307–12 (1998).

116. Chan, D. C., Fass, D., Berger, J. M. & Kim, P. S. Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89,** 263–73 (1997).
117. Lu, M., Blacklow, S. & Kim, P. S. A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nat Struct Biol* **2,** 1075–1082 (1995).
118. Tan, K., Liu, J., Wang, J., Shen, S. & Lu, M. Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 12303–8 (1997).
119. Hinz, A. *et al.* Crystal Structure of the neutralizing Llama VHH D7 and its mode of HIV-1 gp120 interaction. *PLoS One* **5,** 28–32 (2010).
120. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J. & Wiley, D. C. Atomic structure of the ectodomain from HIV-1 gp41. *Nature* **387,** 426–430 (1997).
121. Melikyan, G. B. *et al.* Evidence that the transition of HIV-1 gp41 into a six-helix bundle, not the bundle configuration, induces membrane fusion. *J. Cell Biol.* **151,** 413–23 (2000).
122. Wilen, C. B., Tilton, J. C. & Doms, R. W. HIV: Cell binding and entry. *Cold Spring Harb. Perspect. Med.* **2,** a006866 (2012).
123. Sourisseau, M., Sol-Foulon, N., Porrot, F., Blanchet, F. & Schwartz, O. Inefficient human immunodeficiency virus replication in mobile lymphocytes. *J. Virol.* **81,** 1000–12 (2007).
124. Chen, P., Hübner, W., Spinelli, M. A. & Chen, B. K. Predominant mode of human immunodeficiency virus transfer between T cells is mediated by sustained Env-dependent neutralization-resistant virological synapses. *J. Virol.* **81,** 12582–12595 (2007).
125. Mazurov, D., Ilinskaya, A., Heidecker, G., Lloyd, P. & Derse, D. Quantitative comparison of HTLV-1 and HIV-1 cell-to-cell infection with new replication dependent vectors. *PLoS Pathog.* **6,** e1000788 (2010).
126. Dimitrov, D. S. *et al.* Quantitation of human immunodeficiency virus type 1 infection kinetics. *J. Virol.* **67,** 2182–2190 (1993).
127. Carr, J. M., Hocking, H., Li, P. & Burrell, C. J. Rapid and Efficient Cell-to-Cell Transmission of Human Immunodeficiency Virus Infection from Monocyte-Derived Macrophages to Peripheral Blood Lymphocytes. *Virology* **265,** 319–329 (1999).
128. Sato, H., Orenstein, J., Dimitrov, D. & Martin, M. Cell-to-cell spread of HIV-1 occurs within minutes and may not involve the participation of virus particles. *Virology* **186,** 712–24 (1992).
129. Martin, N. *et al.* Virological synapse-mediated spread of human immunodeficiency virus type 1 between T cells is sensitive to entry inhibition. *J. Virol.* **84,** 3516–3527 (2010).
130. Sattentau, Q. Avoiding the void: cell-to-cell spread of human viruses. *Nat. Rev. Microbiol.* **6,** 815–826 (2008).
131. Jolly, C., Kashefi, K., Hollinshead, M. & Sattentau, Q. J. HIV-1 cell to cell transfer across an Env-induced, actin-dependent synapse. *J. Exp. Med.* **199,** 283–293 (2004).
132. Igakura, T. *et al.* Spread of HTLV-I between lymphocytes by virus-induced polarization of the cytoskeleton. *Science* **299,** 1713–1716 (2003).
133. McDonald, D. *et al.* Recruitment of HIV and its receptors to dendritic cell-T cell junctions. *Science* **300,** 1295–1297 (2003).

134. Jolly, C. & Sattentau, Q. J. Retroviral spread by induction of virological synapses. *Traffic* **5,** 643–650 (2004).

135. Rudnicka, D. *et al.* Simultaneous cell-to-cell transmission of human immunodeficiency virus to multiple targets through polysynapses. *J. Virol.* **83,** 6234–46 (2009).

136. Sowinski, S. *et al.* Membrane nanotubes physically connect T cells over long distances presenting a novel route for HIV-1 transmission. *Nat. Cell Biol.* **10,** 211–219 (2008).

137. Sherer, N. M., Lehmann, M. J., Jimenez-soto, L. F., Pypaert, M. & Mothes, W. Retroviruses can establish filopodial bridges for efficient cell-to-cell transmission. *Nat. Cell Biol.* **9,** 310–315 (2007).

138. Casartelli, N. HIV-1 Cell-to-Cell Transmission and Antiviral Strategies: An Overview. *Curr Drug Targets* **17,** 65–75 (2015).

139. Miyauchi, K., Kim, Y., Latinovic, O., Morozov, V. & Melikyan, G. B. HIV Enters Cells via Endocytosis and Dynamin-Dependent Fusion with Endosomes. *Cell* **137,** 433–444 (2009).

140. de la Vega, M. *et al.* Inhibition of HIV-1 endocytosis allows lipid mixing at the plasma membrane, but not complete fusion. *Retrovirology* **8,** 99 (2011).

141. Daecke, J., Fackler, O. T., Dittmar, M. T. & Krausslich, H.-G. Involvement of Clathrin-Mediated Endocytosis in Human Immunodeficiency Virus Type 1 Entry. *J. Virol.* **79,** 1581–1594 (2005).

142. Sloan, R. D. *et al.* Productive entry of HIV-1 during cell-to-cell transmission via dynamin-dependent endocytosis. *J. Virol.* **87,** 8110–23 (2013).

143. Barre-Sinoussi, F., Ross, A. L. & Delfraissy, J. F. Past, present and future: 30 years of HIV research. *Nat Rev Microbiol* **11,** 877–883 (2013).

144. Saison, J., Cotte, L., Chidiac, C. & Ferry, T. Fatal cumulative toxicities of HAART in a stable, AIDS-free, HIV-infected patient. *BMJ Case Rep.* **2012,** (2012).

145. Warriner, A. H., Burkholder, G. A. & Overton, E. T. HIV-related metabolic comorbidities in the current ART era. *Infect. Dis. Clin. North Am.* **28,** 457–76 (2014).

146. Ruelas, D. S. & Greene, W. C. An integrated overview of HIV-1 latency. *Cell* **155,** 519–29 (2013).

147. Lassen, K., Han, Y., Zhou, Y., Siliciano, J. & Siliciano, R. F. The multifactorial nature of HIV-1 latency. *Trends Mol. Med.* **10,** 525–31 (2004).

148. Siliciano, R. F. & Greene, W. C. HIV latency. *Cold Spring Harb. Perspect. Med.* **1,** a007096 (2011).

149. Coiras, M., López-Huertas, M. R., Pérez-Olmeda, M. & Alcamí, J. Understanding HIV-1 latency provides clues for the eradication of long-term reservoirs. *Nat. Rev. Microbiol.* **7,** 798–812 (2009).

150. Wainberg, M. A., Zaharatos, G. J. & Brenner, B. G. Development of Antiretroviral Drug Resistance. *N. Engl. J. Med.* **365,** 637–646 (2011).

151. Menéndez-Arias, L. & Alvarez, M. Antiretroviral therapy and drug resistance in human immunodeficiency virus type 2 infection. *Antiviral Res.* **102,** 70–86 (2014).

152. Agosto, L. M., Zhong, P., Munro, J. & Mothes, W. Highly active antiretroviral therapies are effective against HIV-1 cell-to-cell transmission. *PLoS Pathog.* **10,** e1003982 (2014).

153. Titanji, B. K., Aasa-Chapman, M., Pillay, D. & Jolly, C. Protease inhibitors effectively block cell-to-cell spread of HIV-1 between T cells. *Retrovirology* **10,** 161 (2013).

154. Permanyer, M. *et al.* Antiretroviral Agents Effectively Block HIV Replication after Cell-to-Cell

Transfer. *J. Virol.* **86,** 8773–8780 (2012).

155. Russell, R. A., Martin, N., Mitar, I., Jones, E. & Sattentau, Q. J. Multiple proviral integration events after virological synapse-mediated HIV-1 spread. *Virology* **443,** 143–9 (2013).
156. Del Portillo, A. *et al.* Multiploid inheritance of HIV-1 during cell-to-cell infection. *J. Virol.* **85,** 7169–76 (2011).
157. Dale, B. M. *et al.* Cell-to-cell transfer of HIV-1 via virological synapses leads to endosomal virion maturation that activates viral membrane fusion. *Cell Host Microbe* **10,** 551–62 (2011).
158. Hübner, W. *et al.* Quantitative 3D video microscopy of HIV transfer across T cell virological synapses. *Science* **323,** 1743–7 (2009).
159. Wild, C., Greenwell, T. & Matthews, T. A synthetic peptide from HIV-1 gp41 is a potent inhibitor of virus-mediated cell-cell fusion. *AIDS Res. Hum. Retroviruses* **9,** 1051–1053 (1993).
160. Kilby, J. M. *et al.* Potent suppression of HIV-1 replication in humans by T-20, a peptide inhibitor of gp41-mediated virus entry. *Nat. Med.* **4,** 1302–7 (1998).
161. Abela, I. a. *et al.* Cell-cell transmission enables HIV-1 to evade inhibition by potent CD4bs directed antibodies. *PLoS Pathog.* **8,** (2012).
162. Rimsky, L. T., Shugars, D. C. & Matthews, T. J. Determinants of human immunodeficiency virus type 1 resistance to gp41-derived inhibitory peptides. *J. Virol.* **72,** 986–93 (1998).
163. Wei, X. *et al.* Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. *Antimicrob. Agents Chemother.* **46,** 1896–1905 (2002).
164. Borrego, P. *et al.* Baseline susceptibility of primary HIV-2 to entry inhibitors. *Antivir. Ther.* **17,** 565–570 (2012).
165. Witvrouw, M. *et al.* Susceptibility of HIV-2, SIV and SHIV to various anti-HIV-1 compounds: implications for treatment and postexposure prophylaxis. *Antivir. Ther.* **9,** 57–65 (2004).
166. Reeves, J. D. *et al.* Primary human immunodeficiency virus type 2 (HIV-2) isolates infect CD4-negative cells via CCR5 and CXCR4: comparison with HIV-1 and simian immunodeficiency virus and relevance to cell tropism in vivo. *J. Virol.* **73,** 7795–804 (1999).
167. Shi, Y. *et al.* Evolution of human immunodeficiency virus type 2 coreceptor usage, autologous neutralization, envelope sequence and glycosylation. *J. Gen. Virol.* **86,** 3385–96 (2005).
168. Pang, W., Tom, S. C. & Zheng, Y. T. Current peptide HIV type-1 fusion inhibitors. *Antivir. Chem. Chemother.* **20,** 1–18 (2009).
169. Woollard, S. M. & Kanmogne, G. D. Maraviroc: a review of its use in HIV infection and beyond. *Drug Des. Devel. Ther.* **9,** 5447–68 (2015).
170. Eggink, D., Berkhout, B. & Sanders, R. W. Inhibition of HIV-1 by fusion inhibitors. *Curr. Pharm. Des.* **16,** 3716–28 (2010).
171. Menéndez-Arias, L. Molecular basis of human immunodeficiency virus type 1 drug resistance: Overview and recent developments. *Antiviral Res.* **98,** 93–120 (2013).
172. Dragic, T. *et al.* A binding pocket for a small molecule inhibitor of HIV-1 entry within the transmembrane helices of CCR5. *Proc. Natl. Acad. Sci.* **97,** 5639–5644 (2000).
173. Fätkenheuer, G. *et al.* Efficacy of short-term monotherapy with maraviroc, a new CCR5 antagonist, in patients infected with HIV-1. *Nat. Med.* **11,** 1170–2 (2005).

174. Cooper, D. A. *et al.* Maraviroc versus efavirenz, both in combination with zidovudine-lamivudine, for the treatment of antiretroviral-naive subjects with CCR5-tropic HIV-1 infection. *J. Infect. Dis.* **201,** 803–13 (2010).

175. Westby, M. *et al.* Emergence of CXCR4-using human immunodeficiency virus type 1 (HIV-1) variants in a minority of HIV-1-infected patients following treatment with the CCR5 antagonist maraviroc is from a pretreatment CXCR4-using virus reservoir. *J. Virol.* **80,** 4909–20 (2006).

176. Caixas, U. *et al.* Long-term maraviroc use as salvage therapy in HIV-2 infection. *J. Antimicrob. Chemother.* **67,** 2538–9 (2012).

177. Armstrong-James, D. *et al.* Clinical outcome in resistant HIV-2 infection treated with raltegravir and maraviroc. *Antiviral Res.* **86,** 224–6 (2010).

178. Massud, I. *et al.* Lack of prophylactic efficacy of oral maraviroc in macaques despite high drug concentrations in rectal tissues. *J. Virol.* **87,** 8952–61 (2013).

179. Neff, C. P., Ndolo, T., Tandon, A., Habu, Y. & Akkina, R. Oral pre-exposure prophylaxis by anti-retrovirals raltegravir and maraviroc protects against HIV-1 vaginal transmission in a humanized mouse model. *PLoS One* **5,** e15257 (2010).

180. Veazey, R. S. *et al.* Protection of rhesus macaques from vaginal infection by vaginally delivered maraviroc, an inhibitor of HIV-1 entry via the CCR5 co-receptor. *J. Infect. Dis.* **202,** 739–44 (2010).

181. Neff, C. P., Kurisu, T., Ndolo, T., Fox, K. & Akkina, R. A topical microbicide gel formulation of CCR5 antagonist maraviroc prevents HIV-1 vaginal transmission in humanized RAG-hu mice. *PLoS One* **6,** e20209 (2011).

182. Dobard, C. W. *et al.* Protection Against Rectal Chimeric Simian/Human Immunodeficiency Virus Transmission in Macaques by Rectal-Specific Gel Formulations of Maraviroc and Tenofovir. *J. Infect. Dis.* **212,** 1988–95 (2015).

183. Mouquet, H. Antibody B cell responses in HIV-1 infection. *Trends Immunol.* **35,** 549–561 (2014).

184. Wei, X. *et al.* Antibody neutralization and escape by HIV-1. *Nature, Publ. online 20 March 2003; | doi10.1038/10.1038/nature01470* **422,** 307 (2003).

185. Richman, D. D., Wrin, T., Little, S. J. & Petropoulos, C. J. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. *Proc. Natl. Acad. Sci.* **100,** 4144–4149 (2003).

186. Simek, M. D. *et al.* Human Immunodeficiency Virus Type 1 Elite Neutralizers: Individuals with Broad and Potent Neutralizing Activity Identified by Using a High-Throughput Neutralization Assay together with an Analytical Selection Algorithm. *J. Virol.* **83,** 7337–7348 (2009).

187. Eroshkin, A. M. *et al.* bNAber: database of broadly neutralizing HIV antibodies. *Nucleic Acids Res.* **42,** D1133-9 (2014).

188. Alam, S. M. *et al.* The Role of Antibody Polyspecificity and Lipid Reactivity in Binding of Broadly Neutralizing Anti-HIV-1 Envelope Human Monoclonal Antibodies 2F5 and 4E10 to Glycoprotein 41 Membrane Proximal Envelope Epitopes. *J. Immunol.* **178,** 4424–4435 (2007).

189. Haynes, B. F. *et al.* Cardiolipin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies. *Science* **308,** 1906–8 (2005).

190. Mouquet, H. *et al.* Memory B Cell Antibodies to HIV-1 gp140 Cloned from Individuals Infected with Clade A and B Viruses. *PLoS One* **6,** e24078 (2011).

191. Mouquet, H. *et al.* Polyreactivity increases the apparent affinity of anti-HIV antibodies by heteroligation. *Nature* **467,** 591–595 (2010).
192. Dennison, S. M. *et al.* Nonneutralizing HIV-1 gp41 Envelope Cluster II Human Monoclonal Antibodies Show Polyreactivity for Binding to Phospholipids and Protein Autoantigens. *J. Virol.* **85,** 1340–1347 (2011).
193. Morris, L. *et al.* Isolation of a Human Anti-HIV gp41 Membrane Proximal Region Neutralizing Antibody by Antigen-Specific Single B Cell Sorting. *PLoS One* **6,** e23532 (2011).
194. Zhu, Z. *et al.* Cross-Reactive HIV-1-Neutralizing Human Monoclonal Antibodies Identified from a Patient with 2F5-Like Antibodies. *J. Virol.* **85,** 11401–11408 (2011).
195. Malbec, M. *et al.* Broadly neutralizing antibodies that inhibit HIV-1 cell to cell transmission. *J. Exp. Med.* **210,** 2813–2821 (2013).
196. Barouch, D. H. *et al.* Protective efficacy of a global HIV-1 mosaic vaccine against heterologous SHIV challenges in rhesus monkeys. *Cell* **155,** 531–9 (2013).
197. Barouch, D. H. *et al.* Therapeutic efficacy of potent neutralizing HIV-1-specific monoclonal antibodies in SHIV-infected rhesus monkeys. *Nature* **503,** 224–8 (2013).
198. Shingai, M. *et al.* Antibody-mediated immunotherapy of macaques chronically infected with SHIV suppresses viraemia. *Nature* **503,** 277–80 (2013).
199. Lai, R. P. J. *et al.* Nef decreases HIV-1 sensitivity to neutralizing antibodies that target the membrane-proximal external region of TMGP41. *PLoS Pathog.* **7,** 1–16 (2011).
200. McCoy, L. E. *et al.* Molecular Evolution of Broadly Neutralizing Llama Antibodies to the CD4-Binding Site of HIV-1. *PLoS Pathog.* **10,** (2014).
201. McCoy, L. E. *et al.* Potent and broad neutralization of HIV-1 by a llama antibody elicited by immunization. *J. Exp. Med.* **209,** 1091–103 (2012).
202. Huang, J. *et al.* Identification of a CD4-Binding-Site Antibody to HIV that Evolved Near-Pan Neutralization Breadth. *Immunity* **45,** 1108–1121 (2016).
203. McCoy, L. E. *et al.* Neutralisation of HIV-1 cell-cell spread by human and llama antibodies. *Retrovirology* **11,** 83 (2014).
204. Duncan, C. J. A. *et al.* High-multiplicity HIV-1 infection and neutralizing antibody evasion mediated by the macrophage-T cell virological synapse. *J. Virol.* **88,** 2025–34 (2014).
205. Klein, F. *et al.* HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. *Nature* **492,** 118–22 (2012).
206. Caskey, M. *et al.* Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117. *Nature* **522,** 487–491 (2015).
207. Lynch, R. M. *et al.* Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection. *Sci. Transl. Med.* **7,** 319ra206 (2015).
208. Shibata, R. *et al.* Neutralizing antibody directed against the HIV-1 envelope glycoprotein can completely block HIV-1/SIV chimeric virus infections of macaque monkeys. *Nat. Med.* **5,** 204–10 (1999).
209. Pietzsch, J. *et al.* A mouse model for HIV-1 entry. *Proc. Natl. Acad. Sci.* **109,** 15859–15864 (2012).
210. Parren, P. W. *et al.* Antibody protects macaques against vaginal challenge with a pathogenic R5

simian/human immunodeficiency virus at serum levels giving complete neutralization in vitro. *J. Virol.* **75,** 8340–7 (2001).

211. Nishimura, Y. *et al.* Determination of a Statistically Valid Neutralization Titer in Plasma That Confers Protection against Simian-Human Immunodeficiency Virus Challenge following Passive Transfer of High-Titered Neutralizing Antibodies. *J. Virol.* **76,** 2123–2130 (2002).
212. Moldt, B. *et al.* Highly potent HIV-specific antibody neutralization in vitro translates into effective protection against mucosal SHIV challenge in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 18921–5 (2012).
213. Mascola, J. R. *et al.* Protection of macaques against vaginal transmission of a pathogenic HIV-1/SIV chimeric virus by passive infusion of neutralizing antibodies. *Nat. Med.* **6,** 207–10 (2000).
214. Hessell, A. J. *et al.* Effective, low-titer antibody protection against low-dose repeated mucosal SHIV challenge in macaques. *Nat. Med.* **15,** 951–4 (2009).
215. Baba, T. W. *et al.* Human neutralizing monoclonal antibodies of the IgG1 subtype protect against mucosal simian-human immunodeficiency virus infection. *Nat. Med.* **6,** 200–6 (2000).
216. Shingai, M. *et al.* Passive transfer of modest titers of potent and broadly neutralizing anti-HIV monoclonal antibodies block SHIV infection in macaques. *J. Exp. Med.* **211,** 2061–74 (2014).
217. Pegu, A. *et al.* Neutralizing antibodies to HIV-1 envelope protect more effectively in vivo than those to the CD4 receptor. *Sci. Transl. Med.* **6,** 243ra88 (2014).
218. Gautam, R. *et al.* A single injection of anti-HIV-1 antibodies protects against repeated SHIV challenges. *Nature* **533,** 105–9 (2016).
219. Ferrari, G. *et al.* Envelope-specific antibodies and antibody-derived molecules for treating and curing HIV infection. *Nat. Rev. Drug Discov.* (2016). doi:10.1038/nrd.2016.173
220. Hale, M. *et al.* Engineering HIV-resistant, anti-HIV chimeric antigen receptor T cells. *Mol. Ther.* **0,**
221. Ali, A. *et al.* HIV-1-Specific Chimeric Antigen Receptors Based on Broadly-Neutralizing Antibodies. *J. Virol.* **90,** JVI.00805-16 (2016).
222. Deeks, S. G. *et al.* A Phase II Randomized Study of HIV-Specific T-Cell Gene Therapy in Subjects with Undetectable Plasma Viremia on Combination Antiretroviral Therapy. *Mol. Ther.* **5,** 788–797 (2002).
223. Brudno, J. N. & Kochenderfer, J. N. Toxicities of chimeric antigen receptor T cells: recognition and management. *Blood* **127,** 3321–3330 (2016).
224. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391,** 806–11 (1998).
225. Elbashir, S. M. *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411,** 494–498 (2001).
226. Agrawal, N. *et al.* RNA interference: biology, mechanism, and applications. *Microbiol. Mol. Biol. Rev.* **67,** 657–85 (2003).
227. Pecot, C., Calin, G., Coleman, R., Lopez-Berestein, G. & Sood, A. RNA interference in the clinic: challenges and future directions. *Nat. Publ. Gr.* **11,** 59–67 (2010).
228. Siomi, H. & Siomi, M. C. Posttranscriptional regulation of microRNA biogenesis in animals. *Mol.*

Cell **38,** 323–32 (2010).

229. Davidson, B. L. & McCray, P. B. Current prospects for RNA interference-based therapies. *Nat. Rev. Genet.* **12,** 329–40 (2011).
230. Daka, A. & Peer, D. RNAi-based nanomedicines for targeted personalized therapy. *Adv. Drug Deliv. Rev.* **64,** 1508–1521 (2012).
231. Liu, Y. P., Schopman, N. C. T. & Berkhout, B. Dicer-independent processing of short hairpin RNAs. *Nucleic Acids Res.* **41,** 3723–33 (2013).
232. Lam, J. K. W., Chow, M. Y. T., Zhang, Y. & Leung, S. W. S. siRNA Versus miRNA as Therapeutics for Gene Silencing. *Mol. Ther. Nucleic Acids* **4,** e252 (2015).
233. Jackson, A. L. *et al.* Expression profiling reveals off-target gene regulation by RNAi. *Nat. Biotechnol.* **21,** 635–7 (2003).
234. Jackson, A. L. & Linsley, P. S. Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nat. Rev. Drug Discov.* **9,** 57–67 (2010).
235. Bitko, V., Musiyenko, A., Shulyayeva, O. & Barik, S. Inhibition of respiratory viruses by nasally administered siRNA. *Nat. Med.* **11,** 50–5 (2005).
236. Soutschek, J. *et al.* Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature* **432,** 173–178 (2004).
237. Braasch, D. A. *et al.* Biodistribution of phosphodiester and phosphorothioate siRNA. *Bioorg. Med. Chem. Lett.* **14,** 1139–43 (2004).
238. Dykxhoorn, D. M. & Lieberman, J. The silent revolution: RNA interference as basic biology, research tool, and therapeutic. *Annu. Rev. Med.* **56,** 401–423 (2005).
239. Akaneya, Y., Jiang, B. & Tsumoto, T. RNAi-induced gene silencing by local electroporation in targeting brain region. *J. Neurophysiol.* **93,** 594–602 (2005).
240. Kishida, T. *et al.* Sequence-specific gene silencing in murine muscle induced by electroporation-mediated transfer of short interfering RNA. *J. Gene Med.* **6,** 105–10 (2004).
241. Tsunoda, S. *et al.* Sonoporation using microbubble BR14 promotes pDNA/siRNA transduction to murine heart. *Biochem. Biophys. Res. Commun.* **336,** 118–27 (2005).
242. Kim, T. W. *et al.* Modification of professional antigen-presenting cells with small interfering RNA in vivo to enhance cancer vaccine potency. *Cancer Res.* **65,** 309–16 (2005).
243. Rossi, J. J. Expression strategies for short hairpin RNA interference triggers. *Hum. Gene Ther.* **19,** 313–7 (2008).
244. Bridge, A. J., Pebernard, S., Ducraux, A., Nicoulaz, A.-L. & Iggo, R. Induction of an interferon response by RNAi vectors in mammalian cells. *Nat. Genet.* **34,** 263–4 (2003).
245. Liu, B. Exploring cell type-specific internalizing antibodies for targeted delivery of siRNA. *Briefings Funct. Genomics Proteomics* **6,** 112–119 (2007).
246. Thomas, M. *et al.* Ligand-targeted delivery of small interfering RNAs to malignant cells and tissues. *Ann. N. Y. Acad. Sci.* **1175,** 32–9 (2009).
247. Dassie, J. P. *et al.* Systemic administration of optimized aptamer-siRNA chimeras promotes regression of PSMA-expressing tumors. *Nat. Biotechnol.* **27,** 839–49 (2009).
248. Zhou, J., Li, H., Li, S., Zaia, J. & Rossi, J. J. Novel Dual Inhibitory Function Aptamer–siRNA

Delivery System for HIV-1 Therapy. *Mol Ther.* **16,** 1481–1489 (2008).

249. McNamara, J. O. *et al.* Cell type-specific delivery of siRNAs with aptamer-siRNA chimeras. *Nat. Biotechnol.* **24,** 1005–1015 (2006).
250. Song, E. *et al.* Antibody mediated in vivo delivery of small interfering RNAs via cell-surface receptors. *Nat. Biotechnol.* **23,** 709–717 (2005).
251. Peer, D., Zhu, P., Carman, C. V, Lieberman, J. & Shimaoka, M. Selective gene silencing in activated leukocytes by targeting siRNAs to the integrin lymphocyte function-associated antigen-1. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 4095–4100 (2007).
252. Yao, Y. *et al.* Targeted delivery of PLK1-siRNA by ScFv suppresses Her2+ breast cancer growth and metastasis. *Sci. Transl. Med.* **4,** 130ra48 (2012).
253. Mo, R. H., Zaro, J. L. & Shen, W.-C. Comparison of cationic and amphipathic cell penetrating peptides for siRNA delivery and efficacy. *Mol. Pharm.* **9,** 299–309 (2012).
254. Kim, W. J. *et al.* Cholesteryl oligoarginine delivering vascular endothelial growth factor siRNA effectively inhibits tumor growth in colon adenocarcinoma. *Mol. Ther.* **14,** 343–50 (2006).
255. Singha, K., Namgung, R. & Kim, W. J. Polymers in small-interfering RNA delivery. *Nucleic Acid Ther.* **21,** 133–47 (2011).
256. Spagnou, S., Miller, A. D. & Keller, M. Lipidic carriers of siRNA: differences in the formulation, cellular uptake, and delivery with plasmid DNA. *Biochemistry* **43,** 13348–56 (2004).
257. Kanasty, R., Dorkin, J. R., Vegas, A. & Anderson, D. Delivery materials for siRNA therapeutics. *Nat. Mater.* **12,** 967–77 (2013).
258. Liu, S., Niger, C., Koh, E. Y. & Stains, J. P. Connexin43 Mediated Delivery of ADAMTS5 Targeting siRNAs from Mesenchymal Stem Cells to Synovial Fibroblasts. *PLoS One* **10,** e0129999 (2015).
259. MacDiarmid, J. A. *et al.* Sequential treatment of drug-resistant tumors with targeted minicells containing siRNA or a cytotoxic drug. *Nat. Biotechnol.* **27,** 643–651 (2009).
260. Dominska, M. & Dykxhoorn, D. M. Breaking down the barriers: siRNA delivery and endosome escape. *J. Cell Sci.* **123,** 1183–9 (2010).
261. Bobbin, M. L., Burnett, J. C. & Rossi, J. J. RNA interference approaches for treatment of HIV-1 infection. *Genome Med.* **7,** 50 (2015).
262. Tyagi, A. *et al.* HIVsirDB: A database of HIV inhibiting siRNAs. *PLoS One* **6,** (2011).
263. DiGiusto, D. L. *et al.* Development of hematopoietic stem cell based gene therapy for HIV-1 infection: considerations for proof of concept studies and translation to standard medical practice. *Viruses* **5,** 2898–919 (2013).
264. Digiusto, D. L. *et al.* RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS- related lymphoma. *Sci. Transl. Med.* **2,** 36ra43 (2011).
265. Wolstein, O. *et al.* Preclinical safety and efficacy of an anti-HIV-1 lentiviral vector containing a short hairpin RNA to CCR5 and the C46 fusion inhibitor. *Mol. Ther. Methods Clin. Dev.* **1,** 11 (2014).
266. Rossi, J. J., June, C. H. & Kohn, D. B. Genetic therapies against HIV. *Nat. Biotechnol.* **25,** 1444–

54 (2007).

267. TERBRAKE, O., KONSTANTINOVA, P., CEYLAN, M. & BERKHOUT, B. Silencing of HIV-1 with RNA Interference: a Multiple shRNA Approach. *Mol. Ther.* **14,** 883–892 (2006).
268. Lee, S.-K. *et al.* Lentiviral delivery of short hairpin RNAs protects CD4 T cells from multiple clades and primary isolates of HIV. *Blood* **106,** 818–26 (2005).
269. ter Brake, O. *et al.* Lentiviral vector design for multiple shRNA expression and durable HIV-1 inhibition. *Mol. Ther.* **16,** 557–64 (2008).
270. Banerjea, A. *et al.* Inhibition of HIV-1 by lentiviral vector-transduced siRNAs in T lymphocytes differentiated in SCID-hu mice and CD34+ progenitor cell-derived macrophages. *Mol. Ther.* **8,** 62–71 (2003).
271. Anderson, J. *et al.* Safety and efficacy of a lentiviral vector containing three anti-HIV genes--CCR5 ribozyme, tat-rev siRNA, and TAR decoy--in SCID-hu mouse-derived T cells. *Mol. Ther.* **15,** 1182–8 (2007).
272. Scherer, L., Rossi, J. J. & Weinberg, M. S. Progress and prospects: RNA-based therapies for treatment of HIV infection. *Gene Ther.* **14,** 1057–64 (2007).
273. Kumar, P. *et al.* T Cell-Specific siRNA Delivery Suppresses HIV-1 Infection in Humanized Mice. *Cell* **134,** 577–586 (2008).
274. Zhou, J. *et al.* Functional in vivo delivery of multiplexed anti-HIV-1 siRNAs via a chemically synthesized aptamer with a sticky bridge. *Mol. Ther.* **21,** 192–200 (2013).
275. Neff, C. P. *et al.* An aptamer-siRNA chimera suppresses HIV-1 viral loads and protects from helper CD4(+) T cell decline in humanized mice. *Sci. Transl. Med.* **3,** 66ra6 (2011).
276. Wheeler, L. A. *et al.* Durable Knockdown and Protection From HIV Transmission in Humanized Mice Treated With Gel-formulated CD4 Aptamer-siRNA Chimeras. *Mol. Ther.* **21,** 1378–1389 (2013).
277. Yang, S., Chen, Y., Ahmadie, R. & Ho, E. A. Advancements in the field of intravaginal siRNA delivery. *J. Control. Release* **167,** 29–39 (2013).
278. Kim, S.-S. *et al.* RNAi-mediated CCR5 silencing by LFA-1-targeted nanoparticles prevents HIV infection in BLT mice. *Mol. Ther.* **18,** 370–6 (2010).
279. Krebs, M. D. & Alsberg, E. Localized, targeted, and sustained siRNA delivery. *Chemistry* **17,** 3054–62 (2011).
280. Pavletich, N. P. & Pabo, C. O. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science* **252,** 809–17 (1991).
281. Wolfe, S. A., Nekludova, L. & Pabo, C. O. DNA recognition by Cys2His2 zinc finger proteins. *Annu. Rev. Biophys. Biomol. Struct.* **29,** 183–212 (2000).
282. Klug, A. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu. Rev. Biochem.* **79,** 213–31 (2010).
283. Pavletich, N. P. & Pabo, C. O. Crystal structure of a five-finger GLI-DNA complex: new perspectives on zinc fingers. *Science* **261,** 1701–7 (1993).
284. Persikov, A. V *et al.* A systematic survey of the Cys2His2 zinc finger DNA-binding landscape. *Nucleic Acids Res.* **43,** 1965–84 (2015).

285. Segal, D. J., Crotty, J. W., Bhakta, M. S., Barbas, C. F. & Horton, N. C. Structure of Aart, a designed six-finger zinc finger peptide, bound to DNA. *J. Mol. Biol.* **363,** 405–21 (2006).

286. Persikov, A. V. & Singh, M. An expanded binding model for Cys2His2 zinc finger protein-DNA interfaces. *Phys Biol.* **8,** 1–23 (2011).

287. Mandell, J. G. & Barbas, C. F. Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases. *Nucleic Acids Res.* **34,** W516-23 (2006).

288. Liu, Q., Segal, D. J., Ghiara, J. B. & Barbas, C. F. Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 5525–5530 (1997).

289. Moore, M., Choo, Y. & Klug, A. Design of polyzinc finger peptides with structured linkers. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 1432–6 (2001).

290. Moore, M., Klug, A. & Choo, Y. Improved DNA binding specificity from polyzinc finger peptides by using strings of two-finger units. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 1437–41 (2001).

291. Kim, J. S. & Pabo, C. O. Getting a handhold on DNA: design of poly-zinc finger proteins with femtomolar dissociation constants. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 2812–7 (1998).

292. Gersbach, C. A., Gaj, T. & Barbas, C. F. Synthetic Zinc Finger Proteins: The Advent of Targeted Gene Regulation and Genome Modification Technologies. *Acc. Chem. Res.* **47,** 2309–18 (2015).

293. Liu, Q., Xia, Z. & Case, C. C. Validated zinc finger protein designs for all 16 GNN DNA triplet targets. *J. Biol. Chem.* **277,** 3850–3856 (2002).

294. Bae, K.-H. *et al.* Human zinc fingers as building blocks in the construction of artificial transcription factors. *Nat. Biotechnol.* **21,** 275–280 (2003).

295. Gaj, T., Liu, J., Anderson, K. E., Sirk, S. J. & Barbas III, C. F. Protein delivery using Cys2-His2 zinc-finger domains. *ACS Chem. Biol.* **Just,** 1–6 (2014).

296. Jamieson, A. C., Miller, J. C. & Pabo, C. O. Drug discovery with engineered zinc-finger proteins. *Nat. Rev. Drug Discov.* **2,** 361–368 (2003).

297. Bhakta, M. S. & Segal, D. J. The generation of zinc finger proteins by modular assembly. *Methods Mol. Biol.* **649,** 3–30 (2010).

298. Gaj, T., Gersbach, C. a. & Barbas, C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31,** 397–405 (2013).

299. Isalan, M., Klug, A. & Choo, Y. A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter. *Nat. Biotechnol.* **19,** 656–60 (2001).

300. Greisman, H. A. A General Strategy for Selecting High-Affinity Zinc Finger Proteins for Diverse DNA Target Sites. *Science (80-.).* **275,** 657–661 (1997).

301. Joung, J. K., Ramm, E. I. & Pabo, C. O. A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 7382–7387 (2000).

302. Maeder, M. L. *et al.* Rapid 'Open-Source' Engineering of Customized Zinc-Finger Nucleases for Highly Efficient Gene Modification. *Mol. Cell* **31,** 294–301 (2008).

303. Wu, H., Yang, W. P. & Barbas, C. F. Building zinc fingers by selection: toward a therapeutic application. *Proc. Natl. Acad. Sci. U. S. A.* **92,** 344–8 (1995).

304. Segal, D. J., Dreier, B., Beerli, R. R. & Barbas, C. F. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target

sequences. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 2758–63 (1999).

305. Dreier, B., Segal, D. J. & Barbas, C. F. Insights into the molecular recognition of the 5'-GNN-3' family of DNA sequences by zinc finger domains. *J. Mol. Biol.* **303,** 489–502 (2000).
306. Dreier, B., Beerli, R. R., Segal, D. J., Flippin, J. D. & Barbas, C. F. Development of Zinc Finger Domains for Recognition of the 5'-ANN-3' Family of DNA Sequences and Their Use in the Construction of Artificial Transcription Factors. *J. Biol. Chem.* **276,** 29466–29478 (2001).
307. Dreier, B. *et al.* Development of zinc finger domains for recognition of the 5'-CNN-3' family DNA sequences and their use in the construction of artificial transcription factors. *J. Biol. Chem.* **280,** 35588–97 (2005).
308. Beerli, R. R. & Barbas, C. F. Engineering polydactyl zinc-finger transcription factors. *Nat. Biotechnol.* **20,** 135–41 (2002).
309. Gonzalez, B. *et al.* Modular System for the Construction of Zinc-Finger Libraries and Proteins. *April* **5,** 791–810 (2010).
310. Beerli, R. R., Segal, D. J., Dreier, B. & Barbas, C. F. Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 14628–33 (1998).
311. Snowden, A. W., Gregory, P. D., Case, C. C. & Pabo, C. O. Gene-specific targeting of H3K9 methylation is sufficient for initiating repression in vivo. *Curr. Biol.* **12,** 2159–2166 (2002).
312. Beerli, R. R., Dreier, B. & Barbas, C. F. Positive and negative regulation of endogenous genes by designed transcription factors. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 1495–500 (2000).
313. Papworth, M. *et al.* Inhibition of herpes simplex virus 1 gene expression by designer zinc-finger transcription factors. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 1621–6 (2003).
314. Rebar, E. J. *et al.* Induction of angiogenesis in a mouse model using engineered transcription factors. *Nat. Med.* **8,** 1427–32 (2002).
315. Laganiere, J. *et al.* An engineered zinc finger protein activator of the endogenous glial cell line-derived neurotrophic factor gene provides functional neuroprotection in a rat model of Parkinson's disease. *J. Neurosci.* **30,** 16469–74 (2010).
316. Jacobsen, F. *et al.* Nucleofection: a new method for cutaneous gene transfer? *J. Biomed. Biotechnol.* **2006,** 26060 (2006).
317. Elouahabi, A. & Ruysschaert, J.-M. Formation and intracellular trafficking of lipoplexes and polyplexes. *Mol. Ther.* **11,** 336–47 (2005).
318. Wang, W., Li, W., Ma, N. & Steinhoff, G. Non-viral gene delivery methods. *Curr. Pharm. Biotechnol.* **14,** 46–60 (2013).
319. Wilber, A. *et al.* A zinc-finger transcriptional activator designed to interact with the gamma-globin gene promoters enhances fetal hemoglobin production in primary human adult erythroblasts. *Blood* **115,** 3033–41 (2010).
320. Tan, S. *et al.* Zinc-finger protein-targeted gene regulation: genomewide single-gene specificity. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 11997–2002 (2003).
321. Thomas, C. E., Ehrhardt, A. & Kay, M. A. Progress and problems with the use of viral vectors for gene therapy. *Nat. Rev. Genet.* **4,** 346–58 (2003).

322. Izmiryan, A., Basmaciogullari, S., Henry, A., Paques, F. & Danos, O. Efficient gene targeting mediated by a lentiviral vector-associated meganuclease. *Nucleic Acids Res.* **39,** 7610–9 (2011).
323. Gaj, T., Guo, J., Kato, Y., Sirk, S. J. & Barbas, C. F. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. *Nat. Methods* **9,** 805–807 (2012).
324. Bailus, B. J. *et al.* Protein Delivery of an Artificial Transcription Factor Restores Widespread Ube3a Expression in an Angelman Syndrome Mouse Brain. *Mol. Ther.* **24,** 548–555 (2016).
325. Liu, J., Gaj, T., Wallen, M. C. & Barbas, C. F. Improved cell-penetrating zinc-finger nuclease proteins for precision genome engineering. *Mol Ther Nucleic Acids* **4,** e232 (2015).
326. Cai, Y., Bak, R. O. & Mikkelsen, J. G. Targeted genome editing by lentiviral protein transduction of zinc-finger and TAL-effector nucleases. *Elife* **3,** e01911 (2014).
327. Bobis-Wozowicz, S. *et al.* Non-integrating gamma-retroviral vectors as a versatile tool for transient zinc-finger nuclease delivery. *Sci. Rep.* **4,** 4656 (2014).
328. He, C. *et al.* Lentiviral protein delivery of meganucleases in human cells mediates gene targeting and alleviates toxicity. *Gene Ther.* **21,** 759–66 (2014).
329. Price, S. A. *et al.* Gene transfer of an engineered transcription factor promoting expression of VEGF-A protects against experimental diabetic neuropathy. *Diabetes* **55,** 1847–1854 (2006).
330. Wang, Z. *et al.* CRISPR/Cas9-Derived Mutations Both Inhibit HIV-1 Replication and Accelerate Viral Escape. *Cell Rep.* **15,** 481–489 (2016).
331. Wang, G., Zhao, N., Berkhout, B. & Das, A. T. CRISPR-Cas9 can inhibit HIV-1 replication but NHEJ repair facilitates virus escape. *Mol. Ther.* **24,** 522 (2016).
332. Eberhardy, S. R. *et al.* Inhibition of Human Immunodeficiency Virus Type 1 Replication with Artificial Transcription Factors Targeting the Highly Conserved Primer-Binding Site Inhibition of Human Immunodeficiency Virus Type 1 Replication with Artificial Transcription Factors Targ. *Society* **80,** 2873–2883 (2006).
333. Segal, D. J. *et al.* Attenuation of HIV-1 Replication in Primary Human Cells with a Designed Zinc Finger Transcription Factor. *J. Biol. Chem.* **279,** 14509–14519 (2004).
334. Margolin, J. F. *et al.* Krüppel-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci. U. S. A.* **91,** 4509–4513 (1994).
335. Urrutia, R. Protein family review KRAB-containing zinc-finger repressor proteins. (2003).
336. Reynolds, L. *et al.* Repression of the HIV-1 5' LTR promoter and inhibition of HIV-1 replication by using engineered zinc-finger transcription factors. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 1615–1620 (2003).
337. Kim, Y. S. *et al.* Artificial zinc finger fusions targeting Sp1-binding sites and the trans-activator-responsive element potently repress transcription and replication of HIV-1. *J. Biol. Chem.* **280,** 21545–21552 (2005).
338. Tebas, P. *et al.* Gene Editing of *CCR5* in Autologous CD4 T Cells of Persons Infected with HIV. *N. Engl. J. Med.* **370,** 901–910 (2014).
339. Perez, E. E. *et al.* Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* **26,** 808–816 (2008).
340. Dando, T. M. & Perry, C. M. Enfuvirtide. *Drugs* **63,** 2755–68 (2003).

341. Wild, C., Oas, T., McDanal, C., Bolognesi, D. & Matthews, T. A synthetic peptide inhibitor of human immunodeficiency virus replication: correlation between solution structure and viral inhibition. *Proc. Natl. Acad. Sci. U. S. A.* **89,** 10537–41 (1992).

342. Matthews, T. J., Wild, C. T., Shugars, D. C., Greenwell, T. K. & Mcdanal, C. B. Peptides corresponding to a predictive a-helical domain of human immunodeficiency virus type 1 gp4l are potent inhibitors of virus infection. **91,** 9770–9774 (1994).

343. Borrego, P. *et al.* An ancestral HIV-2/simian immunodeficiency virus peptide with potent HIV-1 and HIV-2 fusion inhibitor activity. *AIDS* **27,** 1081–90 (2013).

344. Kolkman, J. A. & Law, D. A. Nanobodies - From llamas to therapeutic proteins. *Drug Discov. Today Technol.* **7,** e95–e146 (2010).

345. de Marco, A. Biotechnological applications of recombinant single-domain antibody fragments. *Microb. Cell Fact.* **10,** 44 (2011).

346. Veiga, S., Henriques, S., Santos, N. C. & Castanho, M. Putative role of membranes in the HIV fusion inhibitor enfuvirtide mode of action at the molecular level. *Biochem. J.* **377,** 107–110 (2004).

347. Martins do Canto, A. M. T., Palace Carvalho, A. J., Prates Ramalho, J. P. & Loura, L. M. S. Effect of amphipathic HIV fusion inhibitor peptides on POPC and POPC/Cholesterol membrane properties: A molecular simulation study. *Int. J. Mol. Sci.* **14,** 14724–43 (2013).

348. Gonçalves, J. & Aires da Silva, F. Engineering rabbit antibody variable domains and uses thereof. WO2008136694A1 (2008).

349. Goncalves, J. *et al.* Functional neutralization of HIV-1 Vif protein by intracellular immunization inhibits reverse transcription and viral replication. *J. Biol. Chem.* **277,** 32036–32045 (2002).

350. Oliveira, S. S. *et al.* Assessing combinatorial strategies to multimerize libraries of single-domain antibodies. *Biotechnol. Appl. Biochem.* **59,** 193–204 (2012).

351. Firth, A. E. & Patrick, W. M. GLUE-IT and PEDEL-AA: new programmes for analyzing protein diversity in randomized libraries. *Nucleic Acids Res.* **36,** W281–W285 (2008).

352. Barbas, C. F., Kang, a S., Lerner, R. a & Benkovic, S. J. Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proceedings of the National Academy of Sciences of the United States of America* **88,** 7978–82 (1991).

353. Barbas, C. F. I., Burton, D. R., Scott, J. K. & Silverman, G. J. *Phage display: a laboratory manual*. (Cold Spring Harbor Laboratory Press, 2000).

354. Rato, S. *et al.* Novel HIV-1 knockdown targets identified by an enriched kinases/phosphatases shRNA library using a long-term iterative screen in Jurkat T-cells. *PLoS One* **5,** e9276 (2010).

355. Calado, M. *et al.* Coreceptor usage by HIV-1 and HIV-2 primary isolates: The relevance of CCR8 chemokine receptor as an alternative coreceptor. *Virology* **408,** 174–182 (2010).

356. Schwartz, O., Alizon, M., Heard, J. & Danos, O. Impairment of T cell receptor-dependent stimulation in CD4+ lymphocytes after contact with membrane-bound HIV-1 envelope glycoprotein. *Virology* **198,** 360–5 (1994).

357. Szoka, F. *et al.* Preparation of unilamellar liposomes of intermediate size (0.1-0.2 mumol) by a combination of reverse phase evaporation and extrusion through polycarbonate membranes. *Biochim Biophys Acta* **601,** 559–71 (1980).

358. Ladokhin, A. S., Jayasinghe, S. & White, S. H. How to measure and analyze tryptophan fluorescence in membranes properly, and why bother? *Anal. Biochem.* **285,** 235–45 (2000).

359. Matos, P. M., Franquelim, H. G., Castanho, M. A. R. B. & Santos, N. C. Quantitative assessment of peptide-lipid interactions. Ubiquitous fluorescence methodologies. *Biochim. Biophys. Acta - Biomembr.* **1798,** 1999–2012 (2010).

360. Kabat, E., Wu, T. & Bilowsky, H. *Sequences of immunoglobulin chains*. (National Institutes for Health: NIH Publication 80-2008, 1979).

361. Heidmann, O. & Rougeon, F. Immunoglobulin kappa light-chain diversity in rabbit is based on the 3' length heterogeneity of germ-line variable genes. *Nature* **311,** 74–6 (1984).

362. Birtalan, S., Fisher, R. D. & Sidhu, S. S. The functional capacity of the natural amino acids for molecular recognition. *Mol. Biosyst.* **6,** 1186–1194 (2010).

363. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–1612 (2004).

364. Chan, D. C., Chutkowski, C. T. & Kim, P. S. Evidence that a prominent cavity in the coiled coil of HIV type 1 gp41 is an attractive drug target. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 15613–7 (1998).

365. Peeters, M., Toure-Kane, C. & Nkengasong, J. Genetic diversity of HIV in Africa: impact on diagnosis, treatment, vaccine development and trials. *AIDS* **17,** 2547–2560 (2003).

366. Adachi, a *et al.* Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J. Virol.* **59,** 284–291 (1986).

367. Brügger, B. *et al.* The HIV lipidome: a raft with an unusual composition. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 2641–2646 (2006).

368. Figueira, T. N., Veiga, A. S. & Castanho, M. A. R. B. The interaction of antibodies with lipid membranes unraveled by fluorescence methodologies. *J. Mol. Struct.* **1077,** 114–120 (2014).

369. Söderlind, E., Vergeles, M. & Borrebaeck, C. A. K. Domain libraries: synthetic diversity for de novo design of antibody V-regions. *Gene* **160,** 269–272 (1995).

370. Wesolowski, J. *et al.* Single domain antibodies: promising experimental and therapeutic tools in infection and immunity. *Med. Microbiol. Immunol.* **198,** 157–74 (2009).

371. Lu, J. *et al.* Rapid emergence of enfuvirtide resistance in HIV-1-infected patients: results of a clonal analysis. *J. Acquir. Immune Defic. Syndr.* **43,** 60–4 (2006).

372. Sen, J. *et al.* Alanine scanning mutagenesis of HIV-1 gp41 heptad repeat 1: Insight into the gp120-gp41 interaction. *Biochemistry* **49,** 5057–5065 (2010).

373. Franquelim, H. G. *et al.* Anti-HIV-1 antibodies 2F5 and 4E10 interact differently with lipids to bind their epitopes. *AIDS* **25,** 419–428 (2011).

374. Vauquelin, G. & Packeu, A. Ligands, their receptors and ... plasma membranes. *Mol. Cell. Endocrinol.* **311,** 1–10 (2009).

375. Verkoczy, L. & Diaz, M. Autoreactivity in HIV-1 broadly neutralizing antibodies: implications for their function and induction by vaccination. *Curr. Opin. HIV AIDS* **9,** 224–34 (2014).

376. McCoy, L. E. & Weiss, R. a. Neutralizing antibodies to HIV-1 induced by immunization. *J. Exp. Med.* **210,** 209–23 (2013).

377. Forsman, A. *et al.* Llama antibody fragments with cross-subtype human immunodeficiency virus type 1 (HIV-1)-neutralizing properties and high affinity for HIV-1 gp120. *J. Virol.* **82,** 12069–12081 (2008).

378. Koh, W. W. L. *et al.* Generation of a family-specific phage library of llama single chain antibody fragments that neutralize HIV-1. *J. Biol. Chem.* **285,** 19116–24 (2010).

379. Strokappe, N. *et al.* Llama antibody fragments recognizing various epitopes of the CD4bs neutralize a broad range of HIV-1 subtypes A, B and C. *PLoS One* **7,** e33298 (2012).

380. Morais, M. *et al.* Biodistribution of a 67Ga-labeled anti-TNF VHH single-domain antibody containing a bacterial albumin-binding domain (Zag). *Nucl. Med. Biol.* **41,** 1–5 (2014).

381. Aloia, R. C., Tiant, H., Jensen, F. C., Tian, H. & Jensen, F. C. Lipid composition and fluidity of the human immunodeficiency virus envelope and host cell plasma membranes. *Proc. Natl. Acad. Sci. U. S. A.* **90,** 5181–5185 (1993).

382. Jones, T. D., Crompton, L. J., Carr, F. J. & Baker, M. P. in *Methods in molecular biology* (ed. Clifton, N.) **525,** 405–423, xiv (Humana Press, 2009).

383. McCartney-Francis, N., Skurla, R. M., Mage, R. G. & Bernstein, K. E. Kappa-chain allotypes and isotypes in the rabbit: cDNA sequences of clones encoding b9 suggest an evolutionary pathway and possible role of the interdomain disulfide bond in quantitative allotype expression. *Proc. Natl. Acad. Sci. U. S. A.* **81,** 1794–1798 (1984).

384. Platt, E. J., Bilska, M., Kozak, S. L., Kabat, D. & Montefiori, D. C. Evidence that ecotropic murine leukemia virus contamination in TZM-bl cells does not affect the outcome of neutralizing antibody assays with human immunodeficiency virus type 1. *J. Virol.* **83,** 8289–8292 (2009).

385. Takeuchi, Y., McClure, M. O. & Pizzato, M. Identification of gammaretroviruses constitutively released from cell lines used for human immunodeficiency virus research. *J. Virol.* **82,** 12585–12588 (2008).

386. Derdeyn, C. a *et al.* Sensitivity of human immunodeficiency virus type 1 to the fusion inhibitor T-20 is modulated by coreceptor specificity defined by the V3 loop of gp120. *J. Virol.* **74,** 8358–8367 (2000).

387. Platt, E. J., Wehrly, K., Kuhmann, S. E., Chesebro, B. & Kabat, D. Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1. *J. Virol.* **72,** 2855–2864 (1998).

388. Kimpton, J. & Emerman, M. Detection of replication-competent and pseudotyped human immunodeficiency virus with a sensitive cell line on the basis of activation of an integrated beta-galactosidase gene. *J. Virol.* **66,** 2232–2239 (1992).

389. Weiss, A., Wiskocil, R. L. & Stobo, J. D. The role of T3 surface molecules in the activation of human T cells: a two-stimulus requirement for IL 2 production reflects events occurring at a pre-translational level. *J. Immunol.* **133,** 123–128 (1984).

390. Douglas, N. W., Munro, G. H. & Daniels, R. S. HIV/SIV glycoproteins: structure-function relationships. *J. Mol. Biol.* **273,** 122–149 (1997).

391. Kauffman, K. J., Webber, M. J. & Anderson, D. G. Materials for non-viral intracellular delivery of messenger RNA therapeutics. *J. Control. Release* **240,** 227–234 (2016).

392. Tabernero, J. *et al.* First-in-humans trial of an RNA interference therapeutic targeting VEGF and KSP in cancer patients with liver involvement. *Cancer Discov.* **3,** 406–17 (2013).

393. Schultheis, B. *et al.* First-in-Human Phase I Study of the Liposomal RNA Interference Therapeutic Atu027 in Patients With Advanced Solid Tumors. *J. Clin. Oncol.* **32,** 4141–4148 (2014).

394. Iyer, A. K., Khaled, G., Fang, J. & Maeda, H. Exploiting the enhanced permeability and retention effect for tumor targeting. *Drug Discov. Today* **11,** 812–818 (2006).

395. Noguchi, Y. *et al.* Early phase tumor accumulation of macromolecules: a great difference in clearance rate between tumor and normal tissues. *Jpn. J. Cancer Res.* **89,** 307–14 (1998).

396. Amara, A. *et al.* HIV coreceptor downregulation as antiviral principle: SDF-1alpha-dependent internalization of the chemokine receptor CXCR4 contributes to inhibition of HIV replication. *J. Exp. Med.* **186,** 139–146 (1997).

397. Kularatne, S. A. *et al.* A CXCR4-targeted site-specific antibody-drug conjugate. *Angew. Chemie - Int. Ed.* **53,** 11863–11867 (2014).

398. Egorova, A. *et al.* Chemokine-derived peptides as carriers for gene delivery to CXCR4 expressing cells. *J. Gene Med.* **11,** 772–81 (2009).

399. Egorova, A., Bogacheva, M., Shubina, A., Baranov, V. & Kiselev, A. Development of a receptor-targeted gene delivery system using CXCR4 ligand-conjugated cross-linking peptides. *J. Gene Med.* **16,** 336–51 (2014).

400. Rosen, C. A., Sodroski, J. G., Campbell, K. & Haseltine, W. A. Construction of recombinant murine retroviruses that express the human T-cell leukemia virus type II and human T-cell lymphotropic virus type III trans activator genes. *J. Virol.* **57,** 379–84 (1986).

401. Terwilliger, E., Proulx, J., Sodroski, J. & Haseltine, W. A. Cell lines that express stably env gene products from three strains of HIV-1. *J. Acquir. Immune Defic. Syndr.* **1,** 317–23 (1988).

402. Parent, M. *et al.* Poly(ADP-ribose) polymerase-1 is a negative regulator of HIV-1 transcription through competitive binding to TAR RNA with Tat.positive transcription elongation factor b (p-TEFb) complex. *J. Biol. Chem.* **280,** 448–57 (2005).

403. Smith, S. D. *et al.* Monoclonal antibody and enzymatic profiles of human malignant T-lymphoid cells and derived cell lines. *Cancer Res.* **44,** 5657–60 (1984).

404. Saayman, S. *et al.* The efficacy of generating three independent anti-HIV-1 siRNAs from a single U6 RNA Pol III-Expressed long hairpin RNA. *PLoS One* **3,** (2008).

405. Midelfort, K. S. *et al.* Substantial energetic improvement with minimal structural perturbation in a high affinity mutant antibody. *J. Mol. Biol.* **343,** 685–701 (2004).

406. Jähnichen, S. *et al.* CXCR4 nanobodies (VHH-based single variable domains) potently inhibit chemotaxis and HIV-1 replication and mobilize stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 20565–20570 (2010).

407. Blanchetot, C. & et al. US 2011/0318347 A1. (2011).

408. Cunha-santos, C., Figueira, T. N., Borrego, P. & Oliveira, S. S. Development of Synthetic Light-Chain Antibodies as Novel and Potent HIV Fusion Inhibitors. *Aids* **30,** 1691–701 (2016).

409. Hou, P. *et al.* Genome editing of CXCR4 by CRISPR/cas9 confers cells resistant to HIV-1 infection. *Sci. Rep.* **5,** 15577 (2015).

410. Perdig??o, P., Gaj, T., Santa-Marta, M., Barbas, C. F. & Goncalves, J. Reactivation of latent HIV-1 expression by engineered TALE transcription factors. *PLoS One* **11,** 1–18 (2016).

411. Zuckerman, J. E. & Davis, M. E. Clinical experiences with systemically administered siRNA-based therapeutics in cancer. *Nat. Rev. Drug Discov.* **14,** 843–856 (2015).

412. Lorenzer, C., Dirin, M., Winkler, A.-M., Baumann, V. & Winkler, J. Going beyond the liver: Progress and challenges of targeted delivery of siRNA therapeutics. *J. Control. Release* **203,** 1–15 (2015).

413. Egorova, A. *et al.* CXCR4-targeted modular peptide carriers for efficient anti-VEGF siRNA delivery. *Int. J. Pharm.* **515,** 431–440 (2016).

414. Kumar, P. *et al.* T Cell-Specific siRNA Delivery Suppresses HIV-1 Infection in Humanized Mice. *Cell* **134,** 577–586 (2008).

415. Chen, Z. *et al.* Receptor-mediated delivery of engineered nucleases for genome modification. *Nucleic Acids Res.* **41,** 1–10 (2013).

416. Zhao, X. *et al.* Intracellular delivery of artificial transcription factors fused to the protein transduction domain of HIV-1 Tat. *Protein Expr. Purif.* **90,** 27–33 (2013).

417. Berahovich, R. D., Lai, N. L. N. L., Wei, Z., Lanier, L. L. & Schall, T. J. Evidence for NK cell subsets based on chemokine receptor expression. *J. Immunol.* **177,** 7833–7840 (2006).

418. Poles, M. A., Elliott, J., Taing, P., Anton, P. A. & Chen, I. S. Y. A Preponderance of CCR5+ CXCR4+ Mononuclear Cells Enhances Gastrointestinal Mucosal Susceptibility to Human Immunodeficiency Virus Type 1 Infection. *J. Virol.* **75,** 8390–8399 (2001).

419. Zhou, J. *et al.* Selection, characterization and application of new RNA HIV gp 120 aptamers for facile delivery of Dicer substrate siRNAs into HIV infected cells. *Nucleic Acids Res.* **37,** 3094–3109 (2009).

420. Darbha, R. *et al.* Crystal Structure of the Broadly Cross-Reactive HIV-1-Neutralizing Fab X5 and Fine Mapping of Its Epitope. *Biochemistry* **43,** 1410–1417 (2004).

421. Kwong, P. D. *et al.* Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393,** 648–659 (1998).

422. Huang, C. *et al.* Structure of a V3-containing HIV-1 gp120 core. *Science* **310,** 1025–8 (2005).

423. Pincus, S. H. *et al.* In vivo efficacy of anti-glycoprotein 41, but not anti-glycoprotein 120, immunotoxins in a mouse model of HIV infection. *J. Immunol.* **170,** 2236–41 (2003).

424. Pincus, S. H. *et al.* In vitro effects of anti-HIV immunotoxins directed against multiple epitopes on HIV type 1 envelope glycoprotein 160. *AIDS Res. Hum. Retroviruses* **12,** 1041–51 (1996).

425. Bleul, C. C., Wu, L., Hoxie, J. A., Springer, T. A. & Mackay, C. R. The HIV coreceptors CXCR4 and CCR5 are differentially expressed and regulated on human T lymphocytes. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 1925–30 (1997).

426. Darcis, G., Van Driessche, B. & Van Lint, C. HIV Latency: Should We Shock or Lock? *Trends Immunol.* (2017). doi:10.1016/j.it.2016.12.003